2019

AN OMICS BASED APPROACH FOR THE IDENTIFICATION OF BIOMARKERS IN NON-ALCOHOLIC FATTY LIVER DISEASE USING IN VITRO MODELS OF HEPATIC STEATOSIS

Anitha Saravanakumar
University of Rhode Island, asaravanakumar@uri.edu

AN OMICS BASED APPROACH FOR THE IDENTIFICATION OF BIOMARKERS IN NON-ALCOHOLIC FATTY LIVER DISEASE USING IN VITRO MODELS OF HEPATIC STEATOSIS

BY

ANITHA SARAVANAKUMAR

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

PHARMACEUTICAL SCIENCES

UNIVERSITY OF RHODE ISLAND

2019

DOCTOR OF PHILOSOPHY DISSERTATION

OF

ANITHA SARAVANAKUMAR

APPROVED:

Thesis Committee:

Major Professor Fatemeh Akhlaghi

Angela Slitt

Ingrid Lofgren

Nasser H. Zawia

DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND

2019

ABSTRACT

Nonalcoholic fatty liver disease (NAFLD) is an “umbrella” term for the broad spectrum of the disease that begins from the simple steatosis to more progressive stages of nonalcoholic steatohepatitis (NASH), that includes hepatocellular ballooning, inflammation, and fibrosis. NAFLD is a growing epidemic globally, with 25% of the population predicted to be diagnosed with this disease. Liver biopsy is the only definitive method of diagnosis, despite the widespread use of sonography and elastography to predict the disease state. There is no current FDA approved medication for NAFLD/NASH. This is partly due to the lack of translatable disease model to predict the whole spectrum of the disease in humans as well as lack of definitive biomarker to predict the disease state. The goal of this dissertation is (1) To build an in vitro model for NAFLD to study the influence of fat over-load on xenobiotic as well as lipid metabolizing proteins (2) To use the model to help identify novel biomarkers in liver tissue to characterize the early stages of disease.

Manuscript I: In recent times, there are numerous invitro models relevant to humans developed to predict the disease state. They involve monoculture, coculture and as well as multicellular culture in both 2D and 3D models to best represent the physiology and working of liver in the disease state. In this review, we have explored all the existing in vitro models of NAFLD relevant to humans as well as highlight the technological gaps in the current in vitro models for future development.

Manuscript II: Human hepatic carcinoma cell lines are commonly used in in vitro studies of lipid and xenobiotic metabolism, as well as glucose regulation in normal and disease state. However, their validity is still under debate due to the variable expression of proteins in the cell lines and human hepatocytes. In this present study, we used a data independent acquisition based total protein approach (DIA-TPA) to quantify the protein abundance in

the different cell lines versus (vs.) cryopreserved human hepatocytes (cHH) and human liver tissue (HLT). For this purpose, the global proteome from the whole cell homogenates of HepaRG, HepG2, Huh7 cell lines were compared to cHH and HLT. The MS2 spectra for all detectable peptides were quantified using Spectronaut™. In summary, 2715, 2578, 2874, 2717 and 3083 proteins in, HepaRG, HepG2, Huh7, cHH and HLT, were identified at 1% FDR, respectively. The global proteome of the cHH significantly differed from the cancer hepatic cell lines. Within the cell lines, the global as well as ADME protein profile of HepaRG most closely correlated with cHH, with 84 out of 101 ADME proteins, identified in HepaRG cells. Within gluconeogenesis and glycolysis pathway, Huh7 cell line expressed proteins in high abundance in contrast to the other groups. Therefore, we show that the comparison demonstrates the capability of untargeted global proteomics to detect the differences in protein expression among the different groups. In addition, this study provides a comprehensive database of information to aid study design and model selection.

Manuscript III: To manage NAFLD and related co-morbidities, patients are administered with an array of pharmacological agents. Therefore, understanding of the effect of NAFLD on drug disposition is warranted. Using a HepaRG model, we aimed to mimic steatosis in vitro, and to examine its effects on drug-metabolizing enzymes (DME) and transporters. HepaRG cells, differentiated in-house, were exposed to a mixture of saturated and unsaturated fatty acids (1:2 ratio of 0.5 mM palmitate and oleate complex conjugated to BSA for 72h) and were subjected to RNA sequencing and proteomic analyses. Lipid accumulation was ascertained by Oil Red O (ORO) staining and triglyceride (TG) quantitation and cell viability by WST-1 determination. The treatment condition resulted in ~6-fold increase in TG concentration without reducing cell viability. RNA sequencing of lipid-loaded and control cells identified a total of 393 differentially expressed transcripts

(89 up- and 304 down-regulated). Moreover, lipid loading resulted in significant down-regulation of mRNA transcripts of transcription factors, NR1I2 (-1.18) and HNF4α (-0.55), phase 1 DMEs including CYP1A2 (-3.25), 2B6 (-2.02), 2C8 (-1.48), 2C9 (-2.00),2C19 (-1.32) and 3A4 (-1.87), phase 2 DMEs including UGT1A6 (-0.36) and 2B7 (-1.09), SULT2A1 (-0.75) and 1E1 (-1.41) as well as clinically relevant transporters such as ABCC11 (-1.36), ABCG5 (-1.66), SLC10A1(-1.63) and SLCO2B1(-1.49). However, the protein expression did not show a significant change. Furthermore, lipid loaded cells significantly upregulated AKR1B10 mRNA (2.17) and protein (0.99) that may regulate lipid as well as xenobiotic metabolism.

Manuscript IV: Non-alcoholic fatty liver disease (NAFLD) is a global epidemic, present in over 10% of the world population, despite the majority of the population being undiagnosed. Liver biopsies are the only gold standard available for the confirmation of disease state. Other non-invasive diagnostics such as ultrasound and MRI are either inaccurate or expensive for routine use. Many markers for disease state are available that detect the onset of inflammation and fibrosis from moderate-to-high accuracy. However, there is a huge gap in specific biomarkers that can distinguish the NAFLD liver tissue from normal in early stages of the disease. Using SWATH-MS based Data-independent acquisition (DIA) strategy the dysregulated proteins in the in vitro model of hepatic steatosis was compared with human liver tissue (n=116) showing progressive stages of NAFLD. More than 2,500 proteins were identified in HepaRG and human hepatocyte model as well as human liver tissue. Within the hepatocyte model, 40 proteins were dysregulated in steatosis. These proteins were screened in liver tissue and 6 common proteins were identified. The sensitivity and selectivity of the markers were analyzed using receiver-operative curve (ROC) for the following markers PLIN2 (0.77), ANXA1(0.70),

H2AFY (0.80), SNX1 (0.67), GCHFR (0.69), APO (0.69) and all the above markers showed significance $P < 0.05$.

Conclusion. This work demonstrates that in vitro human relevant disease model has the potential to explain the effect of NAFLD. Subsequently, the in vitro models when used in conjunction with human liver tissue aid in identification of novel biomarkers that may have therapeutic and diagnostic value.

ACKNOWLEDGMENTS

My heartfelt thanks to my major professor, Dr. Fatemeh Akhlaghi, for her continuous support and guidance throughout this dissertation. I am grateful that she provided me an opportunity in her lab and funded me through the entire period of my graduate studies. I am thankful that she gave me an opportunity to choose and pursue the area of my interest and most importantly understand the importance of translatablity in science going from cells to humans. I have learnt so much from her in these past few years and feel better equipped to handle the challenges in both science and life.

I am also grateful to my committee member, Dr. Angela Slitt, for all the collaborations and mentorship. Dr. Ingrid Lofgren and Dr. Matthew Delmonico for mentoring and guiding me in key phases of my PhD. I am grateful for their time, help and advice. I would like to thank Dr. Christopher Hemme, for patiently working with me in trying to understand and interpret informatics data. My sincere thanks to Mrs. Gerralyn Perry and Mrs. Kathy Hayes for all the administrative help and for being so pleasant every time I stopped by. I am thankful to Mrs. Kim Andrews and Dr. Al Bach for providing the support through RI-INBRE Centralized Core Research Facility.

I also would like to acknowledge our Pfizer collaborators, Dr. David Rodrigues, Dr. Jillian Johnson, Dr. Wen He and Dr. Cassandra Tierney from Precision Medicine, Pfizer (Groton, CT). I have a learnt a lot from their expertise and I appreciate their time taken to help us out to run the experiments and execute the data analysis in such a short turn-over period. Most importantly, I appreciate the help and support I received from Dr. Lorenzo Leggio (NIH/NIAAA, Bethesda, MD). The collaborations with NIH and Dr. Akhlaghi provided me the opportunity to learn about clinical trials up close and also gave me the possibility of analyzing the data from different cohorts.

I would also like to take the opportunity to thank Dr. Odette Fahmi and Dr. Ayman El-Kattan for recommending me to Dr. Akhlaghi when I made the transition. I will forever be grateful for all the trust and encouragement given, all through these years.

I would like to thank all my lab mates (Mwlod, Abdullah, Enoch, Sravani, Armin, Rohitash, Ben and Ghadah) for all the collaborative projects we as a team, worked together. I am thankful to the former undergraduate student, Ms. Xin Bush, for helping me culture hepatic cell lines during my maternity break. Also, thankful to our former Pharm.D. student, Ms. Rachel Ryu for her contribution to the comprehensive review on pharmaceutical compounds and for seeing it through till the end.

I am extremely grateful to my mother-and father-in law (Mrs. Lata Arunkumar Dolas and Mr. Arunkumar Krishna Dolas) for being another set of parents to my daughter. Thankful to my brother's family, Mr. Arvind Saravanakumar and Mrs. Ajitha Amarnath and the two adorable nieces (Rhea and Rhinah) for being so accommodative. Most importantly, grateful to my father, Mr. Saravanakumar Appavu, for being so positive and helping me see through everything from a different perspective when I needed it the most.

DEDICATION

This body of work is dedicated to my mother, Mrs. Sundari Saravanakumar, for solely being the strongest pillar of support and guiding force of my life. My husband, Mr. Kunal Arunkumar Dolas, for being my best friend, confidante and for making all the sacrifices through the past few years of our lives. Lastly, my adorable daughter, Taniya, who is the greatest blessing that I could possibly have.

PREFACE

The following body of work is presented in manuscript format which constitute four major chapters. The publication statuses of each manuscript are as follows:

Chapter 1: Preclinical In Vitro Models of Non-Alcoholic Fatty Liver Disease: A Mini-Review

Manuscript 1: In Preparation for submission to Journal of Pharmacology and Experimental Therapeutics

Chapter 2: Comparative Proteomics Analysis of Cryopreserved Human Hepatocytes and Hepatic Cell Lines Using SWATH-MS Reveal Significant Variations in Proteins Involved in Energy, Lipid and Xenobiotic Metabolism

Manuscript 2: In Preparation for submission to Molecular & Cellular Proteomics

Chapter 3: Transcriptomic and Proteomics Analysis of a Lipid-Loaded HepaRG Model for Steatosis Reveals Altered Regulation in Lipid and Xenobiotic Metabolism

Manuscript 3: In Preparation for submission to Liver International

Chapter 4: Data-Independent Acquisition-Based Quantitative Proteomic Analysis Reveals Potential Biomarkers of Non-Alcoholic Fatty Liver Disease

Manuscript 4: In Preparation for submission to Clinical Proteomics

TABLE OF CONTENTS

LIST OF FIGURES

Manuscript 1

Manuscript 2

Manuscript 3

Manuscript 4

LIST OF TABLES

Manuscript 1

Manuscript 2

Manuscript 3

Manuscript 4

JUSTIFICATION OF RESEARCH

1. NAFLD Background

Nonalcoholic fatty liver disease (NAFLD) is a highly prevalent disease in the United States as well as other parts of the world (Bellentani, 2017). A high-fat diet and lack of exercise can lead to the accumulation of free fatty acids and tri-acyl glycerides (TAG) in hepatocytes leading to the development of steatosis (Veteläinen *et al.*, 2007; Koo, 2013). Excess free fatty acids (FA) derived from diet or inflammation-induced lipolysis in adipocytes results in the accumulation of TAGs in the form of lipid droplets (LDs) within the hepatocytes. The occurrence of such LDs greater than 5% of hepatocytes is the diagnostic measure for the onset of NAFLD and liver biopsy is the only form of conclusive testing (Gan, Chitturi and Farrell, 2011; Gorden *et al.*, 2015). The liver steatosis is often termed as the "first-hit" in the progression of NAFLD. The "second-hit" is manifested via the reactive oxygen species (ROS) induced stress leading to the infiltration of inflammatory factors, hepatocellular ballooning and deposition of collagen fibers in hepatocytes as observed in NASH and fibrosis (Caldwell and Argo, 2010). About 29% of the patients diagnosed with NASH develop cirrhosis (Cohen, Horton and Hobbs, 2011). Current treatment involves lifestyle management, which in most cases is proven ineffective. Hence, our long-term goal is to identify non-invasive and conclusive diagnostic methods for the identification of NAFLD as well as to discover a panel of biomarkers that could be potential therapeutic targets for the pharmacotherapy of NAFLD.

2. Importance of studying the disease state with in-vitro systems

To study steatosis, it is crucial to identify the appropriate disease model. Nutrition and energy uptake are the critical determinants for the onset of NAFLD; however, significant differences exist in the manifestation of disease between humans and mice. Diet-induced

mouse models fail to recapitulate the spectrum of NAFLD as seen in humans (Gómez-Lechón, Jover and Donato, 2009; Merrell and Cherrington, 2011). With a high-fat diet (HFD) simple steatosis is established; however, male C57BL/6 mice fail to progress to NASH even in the presence of high fructose-rich diet (Takahashi, Soejima and Fukusato, 2012). To establish NASH, most animal models are fed with methionine-choline-deficient (MCD) diet that results in NASH without inducing obesity. Recent reports, suggest the development of liver steatosis and fibrosis with high fructose (HF) rich diet (Cydylo, Davis and Kavanagh, 2017). In addition, there are several genetically modified mouse models such as, *ob/ob, db/db* that contain mutations in the leptin gene; however, they too fail to recapitulate the full spectrum of the disease (Larter and Yeh, 2008). This physiological mismatch makes it challenging to solely rely upon animal models for investigation of NAFLD.

The use of primary human hepatocytes (PHH) or cryopreserved hepatocytes is considered the gold standard for *in-vitro* studies (Lecluyse and Alexandre, 2010). However, the availability of early-stage biopsy samples, limited time in culture before the onset of dedifferentiation renders the utility of PHH as an ineffective tool for model development. HepaRG cells are a human hepatocellular carcinoma cell line composed of a mixture of both hepatocyte-like and biliary-like cells. They retain hepatic functions, expression of liver-specific genes and drug metabolism capacity at levels comparable to that of PHH (Guillouzo *et al.*, 2007). Furthermore, they preserve their phenotype for over 2 weeks upon differentiation (Aninat *et al.*, 2006). This provides for a good window for model development and treatment modification.

3. The use of 'Omics' based technology in the identification of biomarkers

Omics technology is a broad term to denote the global detection of genes (genomics), mRNA (transcriptomics), protein (proteomics) and metabolites (metabolomics). The total mRNA in a cell or an organism provide a template for the translation of the expressed protein. The transcriptome can be analyzed using RNA sequencing method to provide an unbiased read-out of the all the expressed transcripts, the copy number of the individual transcripts as well as the presence alternatively spliced transcripts between disease and healthy state. This method of untargeted approach aids in the identification of novel genes that are over/under-expressed in disease state and embody a huge potential towards biomarker identification.

Additionally, shotgun or bottom-up proteomic approach is emerging as a powerful tool for analyzing proteolytic peptides of high intensity eluting at the same time in a chromatogram chosen for fragmentation and quantification. The results generated are high throughput and have a complete proteome coverage that out-weighs the traditional labor and resource intensive western blotting techniques (Vildhede *et al.*, 2015; Wiśniewski *et al.*, 2016). Data-dependent acquisition (DDA) methodology has been used more commonly in untargeted analysis (Gu *et al.*, 2018). However, bias associated with the abundant peptides and the lack of sensitivity to quantify low-abundance proteins rendering limitation towards the analysis of individual protein of interest that are present in low expression(Shi *et al.*, 2018). Hence, sequential window acquisition of all theoretical mass spectra (SWATH) analysis (AB Sciex, Concord, Canada) that follows a Data Independent Acquisition (DIA) methodology is used to compare the relative expression pattern among the groups (Huang *et al.*, 2015; Gillet *et al.*, 2012; Gao *et al.*, 2017). This approach is sensitive to capture the low abundant proteins which makes it a unique and reliable tool to capture the differentially regulated proteins in the disease state.

Despite the unbiased, hypothesis generating potential observed in transcriptomics and proteomics methods, there exists a high probability for false positive results. However, high confidence exists when these results from different approach and systems are analyzed together, and similar patterns emerge despite the method of detection. To date, we do not have a unique panel of the biomarkers or a non-invasive method for conclusive diagnosis of NAFLD, and this approach may aid to bridge the gap.

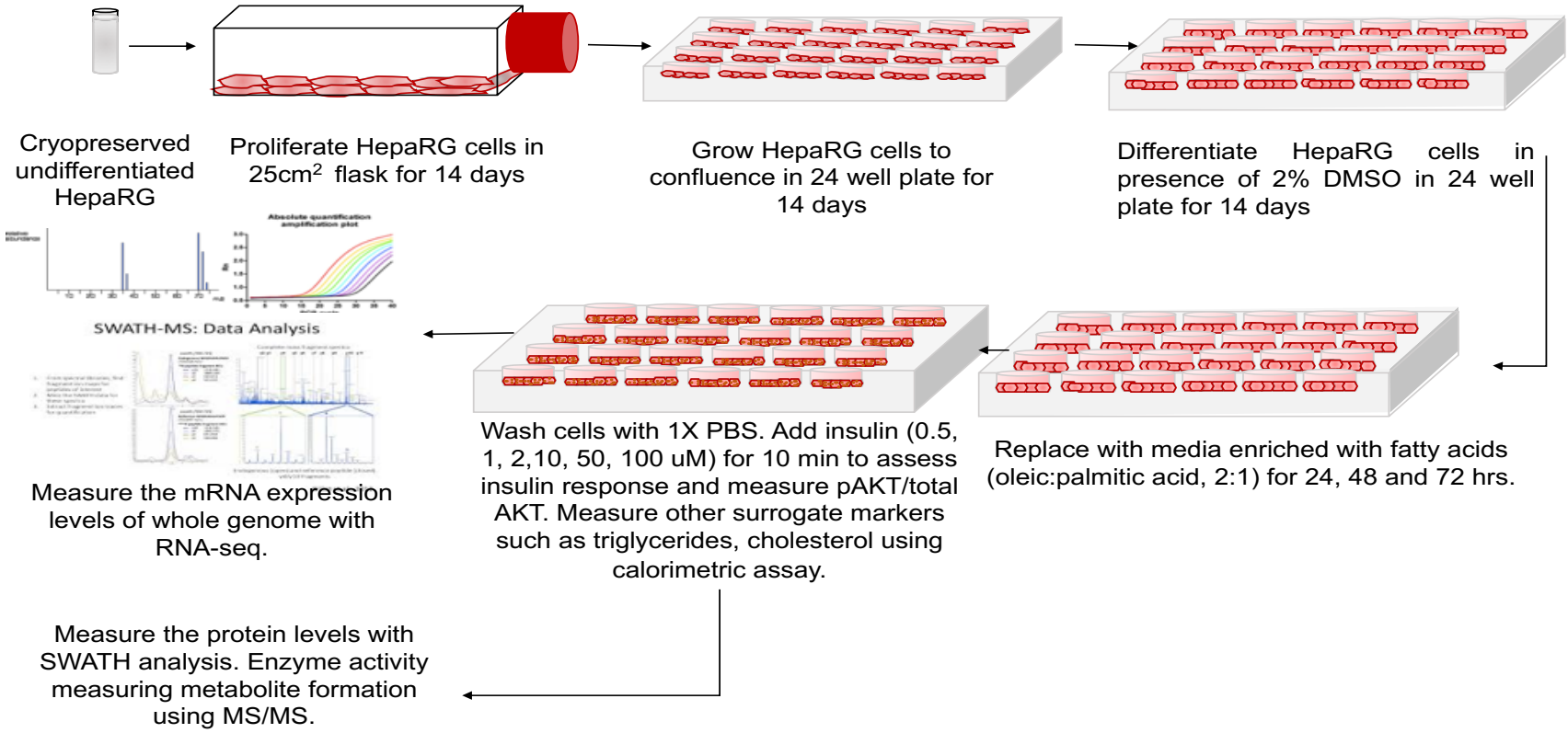

Figure-1 Overview of the invitro model for hepatic steatosis using HepaRG cell line

REFERENCES

Bellentani, S., The epidemiology of non-alcoholic fatty liver disease. *Liver Int* **2017,** *37 Suppl 1*, 81-84.

Veteläinen, R.; van Vliet, A.; Gouma, D. J.; van Gulik, T. M., Steatosis as a risk factor in liver surgery. *Ann Surg* **2007,** *245* (1), 20-30.

Koo, S. H., Nonalcoholic fatty liver disease: molecular mechanisms for the hepatic steatosis. *Clin Mol Hepatol* **2013,** *19* (3), 210-5.

Gan, L.; Chitturi, S.; Farrell, G. C., Mechanisms and implications of age-related changes in the liver: nonalcoholic Fatty liver disease in the elderly. *Curr Gerontol Geriatr Res* **2011,** *2011*, 831536.

Gorden, D. L.; Myers, D. S.; Ivanova, P. T.; Fahy, E.; Maurya, M. R.; Gupta, S.; Min, J.; Spann, N. J.; McDonald, J. G.; Kelly, S. L.; Duan, J.; Sullards, M. C.; Leiker, T. J.; Barkley, R. M.; Quehenberger, O.; Armando, A. M.; Milne, S. B.; Mathews, T. P.; Armstrong, M. D.; Li, C.; Melvin, W. V.; Clements, R. H.; Washington, M. K.; Mendonsa, A. M.; Witztum, J. L.; Guan, Z.; Glass, C. K.; Murphy, R. C.; Dennis, E. A.; Merrill, A. H.; Russell, D. W.; Subramaniam, S.; Brown, H. A., Biomarkers of NAFLD progression: a lipidomics approach to an epidemic. *J Lipid Res* **2015,** *56* (3), 722-36.

Caldwell, S.; Argo, C., The natural history of non-alcoholic fatty liver disease. *Dig Dis* **2010,** *28* (1), 162-8.

Gómez-Lechón, M. J.; Jover, R.; Donato, M. T., Cytochrome p450 and steatosis. *Curr Drug Metab* **2009,** *10* (7), 692-9.

Merrell, M. D.; Cherrington, N. J., Drug metabolism alterations in nonalcoholic fatty liver disease. *Drug Metab Rev* **2011,** *43* (3), 317-34.

Takahashi, Y.; Soejima, Y.; Fukusato, T., Animal models of nonalcoholic fatty liver disease/nonalcoholic steatohepatitis. *World J Gastroenterol* **2012,** *18* (19), 2300-8.

Lecluyse, E. L.; Alexandre, E., Isolation and culture of primary hepatocytes from resected human liver tissue. *Methods Mol Biol* **2010,** *640*, 57-82.

Guillouzo, A.; Corlu, A.; Aninat, C.; Glaise, D.; Morel, F.; Guguen-Guillouzo, C., The human hepatoma HepaRG cells: a highly differentiated model for studies of liver metabolism and toxicity of xenobiotics. *Chem Biol Interact* **2007,** *168* (1), 66-73.

Aninat, C.; Piton, A.; Glaise, D.; Le Charpentier, T.; Langouet, S.; Morel, F.; Guguen-Guillouzo, C.; Guillouzo, A., Expression of cytochromes P450, conjugating enzymes and nuclear receptors in human hepatoma HepaRG cells. *Drug Metab Dispos* **2006,** *34* (1), 75-83.

Vildhede, A.; Wiśniewski, J. R.; Norén, A.; Karlgren, M.; Artursson, P., Comparative Proteomic Analysis of Human Liver Tissue and Isolated Hepatocytes with a Focus on Proteins Determining Drug Exposure. *J Proteome Res* **2015,** *14* (8), 3305-14.

Wiśniewski, J. R.; Vildhede, A.; Norén, A.; Artursson, P., In-depth quantitative analysis and comparison of the human hepatocyte and hepatoma cell line HepG2 proteomes. *J Proteomics* **2016,** *136*, 234-47.

Huang, Q.; Yang, L.; Luo, J.; Guo, L.; Wang, Z.; Yang, X.; Jin, W.; Fang, Y.; Ye, J.; Shan, B.; Zhang, Y., SWATH enables precise label-free quantification on proteome scale. *Proteomics* **2015,** *15* (7), 1215-23.

Gillet, L. C.; Navarro, P.; Tate, S.; Röst, H.; Selevsek, N.; Reiter, L.; Bonner, R.; Aebersold, R., Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics* **2012,** *11* (6), O111.016717.

Gao, Y.; Wang, X.; Sang, Z.; Li, Z.; Liu, F.; Mao, J.; Yan, D.; Zhao, Y.; Wang, H.; Li, P.; Ying, X.; Zhang, X.; He, K., Quantitative proteomics by SWATH-MS reveals sophisticated metabolic reprogramming in hepatocellular carcinoma tissues. *Sci Rep* **2017,** *7*, 45913.

In Preparation for submission to Journal of Pharmacology and Experimental Therapeutics

MANUSCRIPT 1

Preclinical *in vitro* Models of Non-Alcoholic Fatty Liver Disease: A Mini-Review

Anitha Saravanakumar [1] & Fatemeh Akhlaghi [1*]

Affiliations:

[1]*Clinical Pharmacokinetics Research Laboratory, Department of Biomedical and Pharmaceutical Sciences, University of Rhode Island, Kingston, RI, USA*

Corresponding author at:

Fatemeh Akhlaghi, Ph.D.

Clinical Pharmacokinetics Research Laboratory

Department of Biomedical and Pharmaceutical Sciences

The University of Rhode Island, 495A College of Pharmacy, 7 Greenhouse Road,

Kingston, RI 02881, United States.

Email address: fatemeh@uri.edu

Abbreviations

ACC1 - Acetyl CoA carboxylase

ALT - Alanine transaminase

α-SMA - alpha-smooth muscle actin

BPA **-** Bisphenol A

BSEP - Bile Salt Export Pump

CAR - Constitutive Androstane Receptor

CRP - C-reactive protein

CK 8/18 - Cytokeratin 8/18

CYP - Cytochrome P450

COL1A1 - Collagen type 1, alpha 1

DMET - Drug Metabolizing Enzyme and Transporters

FFA - Free fatty acid mixture

FABP-1 - Fatty acid-binding protein-1

FGF21 - Fibroblast growth factor 21

FXR **-** Farnesoid X Receptor

HSCs - Human Stellate Cells

HLO - Human Liver Organoids

IGFβ1 - Insulin-like growth factor 1

IL6 - Interleukin 6

iPSC-Hep - Induced Pluripotent Stem Cells derived Hepatocytes

MPCC - Micropatterned Primary Co-Cultures

MPTC - Micropatterned Primary Tri-Cultures

MRP2 - Multidrug resistance-associated protein 2

NAFL - Nonalcoholic fatty liver disease

NAFLD - Nonalcoholic fatty liver disease

NASH - Nonalcoholic steatohepatitis

NPC - Non-Parenchymal Cells

NTCP - Sodium taurocholate cotransporter protein

OA - Oleic Acid

OSA - Obstructive Sleep Apnea

PA- Palmitic Acid

PCLS - Precision-Cut Liver Slices

PCOS - Polycystic ovarian syndrome

PDK4 - Pyruvate dehydrogenase kinase

PNPLA3 - Patatin-like phospholipase domain-containing protein 3

PHH - Primary Human Hepatocytes

PXR - Pregnane X Receptor

SCD1 b - Stearoyl-CoA desaturase

SREBP - Sterol regulatory element-binding proteins

STC2 - Stanniocalcin-2

TG - Thapsigargin

TAG -Triacylglycerides

TGFβ - Transforming growth factor-beta

TIMP1 - TIMP metallopeptidase inhibitor 1

TM6SF2 - Transmembrane 6 Superfamily Members 2

T2DM - Type 2 Diabetes Mellitus

VLDL - Very-low-density lipoprotein

3D-hLMT - 3D-Human Liver Micro-Tissues

Abstract

Nonalcoholic fatty liver disease (NAFLD) is an “umbrella” term for the broad spectrum of the disease that begins from the simple steatosis to more progressive stages of nonalcoholic steatohepatitis (NASH), that includes hepatocellular ballooning, inflammation, and fibrosis. NAFLD is a growing epidemic globally, with 25% of the population predicted to be diagnosed with this disease. Liver biopsy is the only definitive method of diagnosis, despite the widespread use of sonography and elastography to predict the disease state. There is no current FDA approved medication for NAFLD/NASH, due to the lack of definitive biomarker to predict the disease state or translatable disease and/or model to predict the whole spectrum of the disease in humans. In this review, we have explored all the existing in vitro models of NAFLD relevant to humans as well as highlight the technological gaps in the current in vitro models for future development.

Keywords

Preclinical; NAFLD; in vitro models; NASH; Hepatic steatosis.

1. Introduction

Hepatic steatosis is characterized by the accumulation of lipids (> 5% of the hepatocytes) in the cytoplasmic region as micro and macrovesicular vacuoles. This may progress to nonalcoholic steatohepatitis (NASH), which results in hepatocellular injury (1). Steatosis is often reversible and self-limiting. However, ~25% of the cases progress to NASH, wherein the liver may become fibrotic resulting in permanent damage. In terms of the worldwide incidence, it is noted that extreme variations occur in the reported range between the studies (2). The current estimates around the incidence of nonalcoholic fatty liver (NAFLD), ranges between 20-30% in Western countries and 5-18% in Asia, suggesting it to become the leading cause of end-stage liver disease and transplantation, in the next decade (3). The fluctuations in the estimates, is in part, associated with the lack of definitive biomarkers to characterize the disease state as well as the limitations in current testing methods. Due to these challenges, there are currently no US Food and Drug Administration (FDA)-approved medication to treat NAFLD or NASH (4, 5).

Numerous risk factors have accentuated the development of NAFLD, namely, diet, lifestyle, age, gender, ethnicity, genetics, and metabolic disorders that impact the lipid metabolism and insulin resistance **(Fig. 1.1)** (2). Other risk factors involve, drug-induced toxicity, as seen in tetracycline and amiodarone, and environmental pollutants, such as bisphenol a (BPA), that may induce and ameliorate NAFLD (6, 7). Besides, the metabolic deregulation observed in the disease state may also impact the disposition and dosing of other concomitant medication within the patient population.

To investigate the drug deposition of the administered drugs due to changes in hepatic metabolism in disease state as well as inspect the chemical-induced toxicity that ameliorates NAFLD; early-stage screening tools that most represent the disease state in humans, are warranted. These models are also advantageous during drug development

for examining various molecular entities against therapeutic targets that contribute to disease progression, e.g., FXR agonists (8). Due to the widely acknowledged ethical and financial requirements that limit the animal use in exploratory studies (9), system-oriented approaches involving integrated into in vitro and in silico models to best predict the changes in human liver physiology, are emerging as valuable tools to study the molecular changes involved in the disease state (10).

In the forthcoming sections, we will provide context for the pathology of NAFLD, risk factors associated with the disease progression as well as the current sophisticated in-vitro human models that reflect the different stages underlying the etiology of NAFLD. The aim of this review is to provide comprehensive information on the existing in vitro models of NAFLD to aid the choice of appropriate model as well as highlight the technological gaps in the current in vitro models for future development.

2. Pathology and Risk factors

Lipid accumulation in the hepatocytes is the starting point for hepatic steatosis. However, with the severity of progression, the affected parenchyma is subdivided into three categories: 5%-33%, 34%-66%, and > 66% as mild, moderate, and severe, respectively (11). The occurrence of steatosis is spatially heterogenous in the liver, with lipids localized to zone 3 in the early stages of disease in contrast to panacinar localization, as seen with the progression of steatosis (12). This heterogeneity may lead to occasional overestimation of the degree of steatosis by histopathologists. In addition, conventional imaging (ultrasound, magnetic resonance imaging, or computed tomography), lack the sensitivity to detect hepatic steatosis less than 30% (13). However, newer technologies involving magnetic resonance imaging-estimated proton density fat fraction, and 1H-magnetic resonance spectroscopy detect steatosis with high accuracy (14, 15).

Steatosis often triggers the infiltration of inflammatory modulators via Kupffer cells that result in hepatocyte ballooning, inflammation, and injury. At this point, the disease progresses to nonalcoholic steatohepatitis (NASH), which may lead to fibrosis wherein the healthy cells are replaced with scarred tissue due to the deposition of collagen. Recent findings highlight the promising role of cytokeratin CK 8/18 to be used as a biomarker to detect hepatocyte ballooning as their expression is significantly decreased in comparison to normal hepatocytes (12), as observed in histological staining. Also, CK-18 fragments detected in the plasma positively correlate with the severity of NASH ($p < 0.001$) (16).

NAFLD is observed at a higher incidence in cohorts with metabolic syndrome and cardiovascular risk **(Fig. 1.1)**(17). Additionally, type 2 diabetes mellitus (T2DM) patients display a high risk for developing NASH and a 2-to-4-fold risk of developing NAFL. Ethnicity also plays a critical role with a higher rate of incidence seen in Hispanics and Asians, followed by Caucasians (18). Possession of homozygous *PNPLA3* allele (rs738409) in the Hispanic population contributes to 2-fold higher hepatic lipid accumulation (19). On the other hand, a lower incidence of steatosis is observed in African-Americans in comparison to Caucasians. In addition, genetic polymorphism in the *TM6SF2* gene, involved in VLDL secretion, leads to a higher incidence of NAFLD. The variant frequencies in *TM6SF2* gene is found to be higher in whites, African-American, and Hispanic races (19).

The occurrence of NAFLD increases with age an incidence of 20% in age group under 20 to 40% or more in the age group above 60 (20). Diet enriched with macronutrients and carbohydrates posit high risk in developing NAFL as well as its progression. It can be worsened when coupled with a sedentary lifestyle and smoking (21, 22). Moreover, endocrine dysfunction, such as polycystic ovarian disorder (PCOS), is often characterized by obese and insulin-resistant women (23). There are evidences that

show PCOS with insulin resistance worsens hyperandrogenemia, which leads to NAFL. A striking association of NAFL is also observed in patients with obstructive sleep apnea (OSA), which is caused by a complete or partial obstruction of the airway. The incomplete exchange of gases (hypoxemia and hypercapnia) leads to oxidative stress, recruitment of pro-inflammatory factors, insulin resistance, and endothelial dysfunction. This may lead to the evolution of NAFL to NASH (24).

3. In vitro models of Nonalcoholic fatty liver disease (NAFLD)

The current in vitro models for NAFLD is summarized in **Table 1.1** and the overview of all the models are illustrated in **Figure 1.2**

3.1 Monoculture systems

3.1.1 Hepatocellular Carcinoma (HepG2, Huh7, and HepaRG)

Human hepatocarcinoma cell lines have been widely used in the literature to model NAFLD in vitro. The ease of propagation, reproducibility, and low cost for procurement makes them an ideal candidate for preliminary model development. The majority of the models for steatosis have been developed by exposing the cells to free fatty acids, namely oleate (OA) and palmitate (PA) **(Table 1.1)**. The choice of these fatty acids stems from the high content of oleate and palmitate in triglycerides. Besides, clinical data show varying levels of these free fatty acids (FFA) in blood ranging between ~0.25-3.0 mmol/L, moreover, they are chronically higher in individuals with obesity and/or diabetes.

Acute treatment (24h) with FFA mixture, induced all three cancer cell lines to show dose-dependent lipid and triglyceride accumulation; on the other hand, with palmitate alone, induced apoptosis, oxidative stress, and insulin resistance. Furthermore, the addition of stearic acid showed induction of CYP2E1 activity in addition to the development

of steatosis (26). Apart from fatty acid stimuli, other models assess the impact of glucose, fructose, and insulin on the cell lines to mimic the onset of NAFL, as seen in increased intake of dietary macronutrients (27). Lastly, steatosis was assessed in the HepaRG cell line stimulated with oleate and palmitate, along with varying concentrations of IL6 (28). The presence of IL6 not only induced NAFLD and NASH like phenotype, but also resulted in significant downregulation of CYP2A6, 2B6, 2C19, 2C9, 3A4, and 3A5 mRNA after a 24h treatment (28).

In context to toxicological assessments, Tolosa et al. (2016) demonstrate the ability of HepaRG cell line to appropriately flag the inducers of steatosis upon exposure to 16-known steatogenic compounds (29). Furthermore, treatment with tetracycline and amiodarone after acute and repeated treatments induced steatosis, highlighting the robustness of HepaRG, in predicting the drug-induced hepatotoxicity as well as its utility for understanding the molecular mechanisms involved in the onset and progression of disease state (7). In summary, these reports represent the possible impact of various stimuli on the onset of hepatic steatosis and/or NASH. However, the results from cancer cell lines are viewed with skepticism as the translatability of the observed impact on human hepatocytes is unknown. In addition, it is well-known that Huh7 and HepG2 cell line expresses the drug-metabolizing enzymes in low levels in comparison to human hepatocytes (30, 31). Hence understanding the DME modulation could be a challenge. However, HepaRG cells can be used to overcome this drawback as evidence show the presence of comparable levels of DME to human hepatocytes.

3.1.2 Primary Human Hepatocytes (PHH)

Primary human hepatocytes are the current gold standard to study metabolism and toxicity related effects of drugs. Donato et al. (2006), showed a model using hepatocytes isolated from 17 steatotic and non-steatotic human liver. Upon harvest and seeding, the

cells were treated with oleate and palmitate to maintain the hepatocytes in the state of steatosis for 14 h and 36 h (32). As a result, lipid accumulation and triglyceride increase were observed in the hepatocytes with steatosis. In addition, *CYP1A2, CYP2A6*, *CYP2C9, CYP2D6, CYP2E1, CYP2B6,* and *CYP3A4* mRNA, along with CYP1A2, CYP2A6, CYP2C9, CYP2D6, CYP2E1, and CYP3A4 enzyme activity showed a significant reduction in steatosis (32). This was the first study in which several human P450 enzymes have been comparatively examined in a model of induced steatosis. This closely represented the disease state; however, the hepatocytes in culture, start to lose their metabolic capacity after 72h, and hence, a static, two dimensional (2D)-model cannot be representative of long-term impact.

To mitigate the drawbacks of the 2D static model and to study the long-term impact of triglyceride accumulation, three dimensional (3D) spheroids or 3D models using scaffolds were developed (33). The 3D models using scaffolds were cultured in the microfluidic device and LiverChip platform with continuous perfusion of media to mimic the dynamic sinusoidal flow in the liver (34). This stabilized the hepatic metabolizing capacity to over 14 days in culture. Subsequently, the model was treated with control and FFA mixture (oleate and palmitate, 2:1) for 7d. As a result, increased levels of secreted adipokines were noted; *IGFβ1*, *PDK4*, and *FABP-1* genes associated with lipid metabolism were upregulated; and a decrease in the metabolic rate of CYP3A4 and CYP2C9 were also reported.

3.1.3 Induced Pluripotent Stem Cells derived Hepatocytes (iPSC-Hep)

Despite the fact that human hepatocytes are considered as gold-standard, they represent several limitations, such as donor to donor variability, lack of availability, cost of procurement, variability in proliferation, which introduce mutations and polymorphisms in metabolic markers. Human-induced pluripotent stem cell (hiPSC)-derived hepatocytes

(hiPSC-Hep) generated from mature hepatocytes overcome the limitations observed with PHH. Henceforth, a model using hiPSC-Hep was co-treated with FFAs (oleate and palmitate) and the ER stressor, thapsigargin (TG), to recapitulate NASH (8). This model showed intracellular lipid accumulation via activation of *denovo lipogenesis* as well as its association with ER related UPR network. In addition to the genes involved in lipid metabolisms such as *SREBP-1c, ACC1*, and *SCD1*, and known markers of NASH such as *STC2* and *FGF21* were upregulated. Henceforth, this model showed the potential to induce NAFL and NASH successfully. Additionally, hiPSC-Hep demonstrates an added advantage of having the potential to derive genotype-specific hepatocytes from patient samples to understand the relevance of polymorphic genes (*PNPLA3*) in the signaling network.

3.2 *Co-culture systems*

3.2.1 Micropatterned Primary Co-Cultures (MPCCs)

To maintain the longevity of primary human hepatocyte in culture, MPCCs were developed to study the effects of hyper and hypoglycemia in hepatocytes. MPCCs recapitulate the homotypic interactions with the primary human hepatocytes as well as the heterotypic interactions of hepatocytes with supportive non-parenchymal cells (NPCs) such as fibroblasts (35). MPCC model was facilitated by patterning 24-well plated with rat-tail collagen 1 at 500 μm diameter spaced 1200 μm away from each other. This enabled the selective attachment of hepatocytes to collagen islands along with the 3T3-J2 fibroblasts covering the spaces between the islands. This study, for the first time, showed the culture of PHH with fibroblasts, stabilized the PHH in culture for ~3 weeks, and was used to understand the response to chronic treatment with varying concentrations of glucose (35). Upon treatment with high glucose, insulin-responsive genes *SREBF1 and FASN* were upregulated along with increased lipid accumulation in the hepatocytes,

suggesting the shuttling of glucose to insulin-stimulated lipogenesis in the cells. Also, hyperglycemia downregulated *CYP2E1* transcripts. This model showed 'selective' insulin resistance akin to type 2 diabetes mellitus (T2DM).

3.2.2 MPCCs with activated stellate cells

MPCCs were constructed with activated or growth arrested human stellate cells (HSCs) as the supportive cell type (1 HSC: 12 PHHs using 2.5k HSCs) along with human hepatocytes (36). The addition of proliferative stellate cells resulted in an increase in albumin production and CYP3A4 activity in hepatocytes. However, severe distortion of hepatocyte islands occurred in 2 weeks. Showing that the activated HSCs, proliferative or growth-arrested, cannot support PHH functions to the same extent of 3T3-J2 fibroblasts.

3.3 *Tri-culture systems*

3.3.1 Micropatterned Tri-Cultures (MPTCs)

To stabilize the hepatocytes in long-term culture as well as induce NASH-like phenotype with activated stellate cells, a micropatterned tri-culture model was introduced. In this model, growth-arrested 3T3-J2s (90k cells per 24-well format) were mixed with activated HSCs (2.5k), and this mixture was then seeded onto to micropatterned PHH colonies (30k cells) (36). The ratio of stellate cells to PHH was corresponding to the proportion in the human liver. In such micropatterned tri-cultures (MPTCs), albumin secretion was higher than the levels seen in MPCCs. Besides, the supporting non-parenchymal cells (NPCs) stabilized the hepatocytes in culture. Upon the stimulus from activated stellate cells for 13 d, significant levels of lipid-filled vesicles (containing acyl-glycerols and cholesterol esters) were observed. Besides, the upregulation of IL6 and CRP protein, as seen in NASH, was observed. Downregulation of significant liver-specific genes was observed, such as *PXR, CAR, CYP3A4, CYP2A6, NTCP, MRP2, BSEP, and FXR.* Interestingly, MPTCs show collagen-I deposition by α-SMA-positive HSCs along

with the upregulation in the mRNA of *LOX* and *COL1A1* (36). In summary, MPTCs show NASH/fibrosis like phenotype. Despite the well-maintained architecture, the model occasionally induced distorted hepatocyte islands due to donor-to-donor variability. Henceforth, human hepatocytes, and stellate cell lots in the development of this model must be chosen after careful screening.

3.3.2 Trans-well culture

An in vitro model with the combination of coculture of stellate cells and fibroblasts in trans-well and monoculture on hepatocytes in the culture plates was developed to model NASH. The separation of stellate cells and fibroblasts from hepatocytes helped in understanding the specific impact on hepatocytes. In addition to the culture, the model was engineered to incorporate hemodynamic stress, hepatic sinusoidal flow, transport, and finally exposed to lipotoxic risk factors such as glucose, insulin, free fatty acids (oleate and palmitate) (37). Upon the stimulus, hepatocyte glucose output increased with decreased insulin sensitivity. Increased levels of IL6, IL8, and ALT were observed along with the activation of fibrogenic markers such as TGF-ß. This model showed transcriptomics, proteomics, and lipidomics signature similar to that of NASH liver biopsies.

3.4 *Multi cell-culture systems*

3.4.1 3D-Human Liver Micro-Tissues (3D-hLMT)

3D-hLMT model is the known closest in vitro culture system that mimic the in vivo liver physiology. The model can recapitulate the multicellularity in 3D format and maintain individual cell phenotype for over 5 weeks (38). PHH were co-cultured with NPCs such as hepatic stellate cells, Kupffer cells, and endothelial cells. In this model, 0.5 mM of palmitic acid (PA) induces a robust inflammatory as well as fibro-genic response in the 3D-hLMT. PA significantly induced the expression of collagen related genes*, α-SMA, TIMP1, IL8*

transcripts, as well as *TGFβ* gene activation pathway (38). Nevertheless, owing to its complexity, these models are time-consuming and expensive to establish; in addition, scalability is a limiting factor to use these models in routine high-throughput screening.

3.4.2 Induced Pluripotent Stem Cell (iPSCs) derived organoids

iPSCs derived from 11 human donors (healthy and diseased) were developed to derive organoid comprising of hepatocyte, stellate, and Kupffer and biliary-like cells whose genetic signature was confirmed from transcriptomics analysis. Upon stimulus with fatty acids (800 µm of oleate) for 5 days, the human liver organoids (HLO) showed the stepwise progression of NAFLD from steatosis, ballooning, inflammation, and fibrosis phenotypes recapitulating NASH in the latter stages (39). In addition, atomic force microscopy confirmed the severity of fibrosis conferred in the HLOs. This novel approach sets a unique benchmark to understand the genotype-specific contributions on the progression of disease state as well as contribute to therapies based on personalized treatment. However, the significant limitations associated with this model must also be taken to account. The donor to donor variability might affect the HLOs phenotype, such as NASH and fibrosis (39). Additionally, the time and cost involved in the organoid development may pose further limitations in routine usage.

3.5 *The organotypic ex-vivo culture system*

3.5.1 Precision-Cut Liver Slices (PCLS)

Many advancements have been noticed in the use of 2D and 3D monoculture to mimic NAFLD. However, the complex interactions of multiple cell types such as hepatocytes, Kupffer cells, stellate, and endothelial cells that induce inflammation, necrosis, and consequently, fibrosis in NAFLD, is impossible to be recapitulated in a 2D and 3D-monoculture. Human precision-cut liver slices (hPCL) retain the liver architecture with the multi-cell interactions among the specialized liver-specific cell types as well as infiltrating

lymphocytes to up to 5 days in culture (40). Henceforth, they make an excellent model to study the changes observed during the different stages of NAFLD and test new therapies.

hPCLS model for NAFLD is developed as follows: (1) Resected liver tissue flushed with ice-cold University of Wisconsin (UW) solution and ~ 1 cm thick tissue is prepared for further processing. (2) Following this, a hollow cylindrical tissue drill is used to prepare chunks of 5 mm diameter cores. (3) The cores are inserted in the tissue holder and cut with the microtome (Krumdieck Tissue Slicer). (4) Liver sections of 5 mg weight and ~ 250 µm thick are prepared for further treatments. (5) hPCLS are treated with free fatty acids (a mixture of oleic and linoleic acids, 0.1 mM) to mimic the hepatic steatosis, lipid deposition, and lipotoxicity. Alternatively, opportunities to prepare hPCLS from cirrhotic patients to assess therapeutic interventions also exist for future development. PCLS makes ideal models, especially by maintaining the intact morphology of hepatocytes. However, the limited time in culture and the inability to adapt to genetic changes such as transfection or RNA silencing, makes it challenging to understand the impact of a specific gene in the signaling pathways leading to disease progression.

4. Conclusion

A translatable human in vitro model of NAFLD is the need of the hour. However, a model that fits all-purpose, to study a "multiple-hit" disease state, such as NAFLD, poses a highly relevant challenge to the research community. The majority of models developed thus far represent 2D-static mode involving either PHH or human carcinoma cell lines. However, it is impossible to study the whole spectrum of disease state using such simple systems. Hepatic steatosis was well recapitulated using various stressors, such as free fatty acids, glucose, and fructose, insulin, interleukins as well as other chemicals. A few studies have advanced further, by including microfluidic the model to mimic the sinusoidal flow as well as micropatterning to recapitulate the spatial organization of hepatocytes with

the supporting NPCs. Sophisticated models with 3D-hLMT as well as iPSC-derived organoids, only show a sneak-peak of the innovations coming up soon. This may allow for the whole body to be modeled on the chip to understand the disease state beyond the liver, by including the gut axis as well as adipose tissue. Such innovative tools will contribute to new therapies as well as understanding other complexities that come with the disease.

5. Figures and Tables

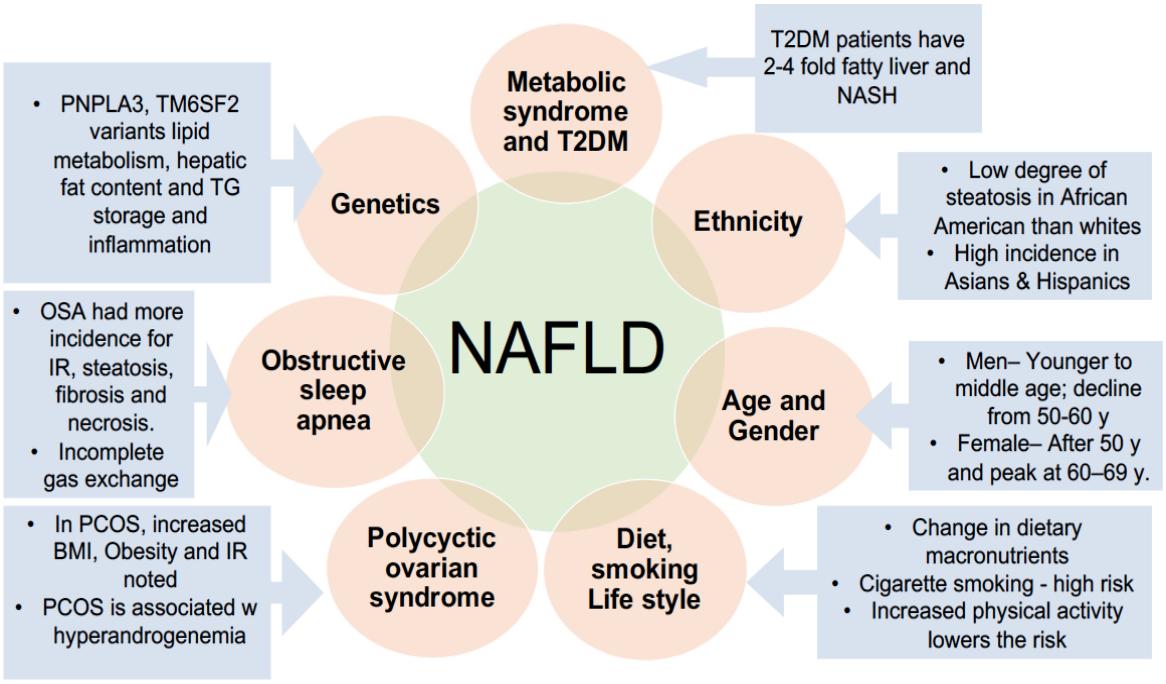

Figure 1. 1: Risk factors associated with nonalcoholic fatty liver disease.

Abbreviations. NAFLD: Nonalcoholic fatty liver disease; NASH: Nonalcoholic steatohepatitis; T2DM: Type 2 diabetes mellitus; PCOS: Polycystic ovarian syndrome; IR: Insulin resistance; OSA: Obstructive sleep apnea; TG: Triglycerides.

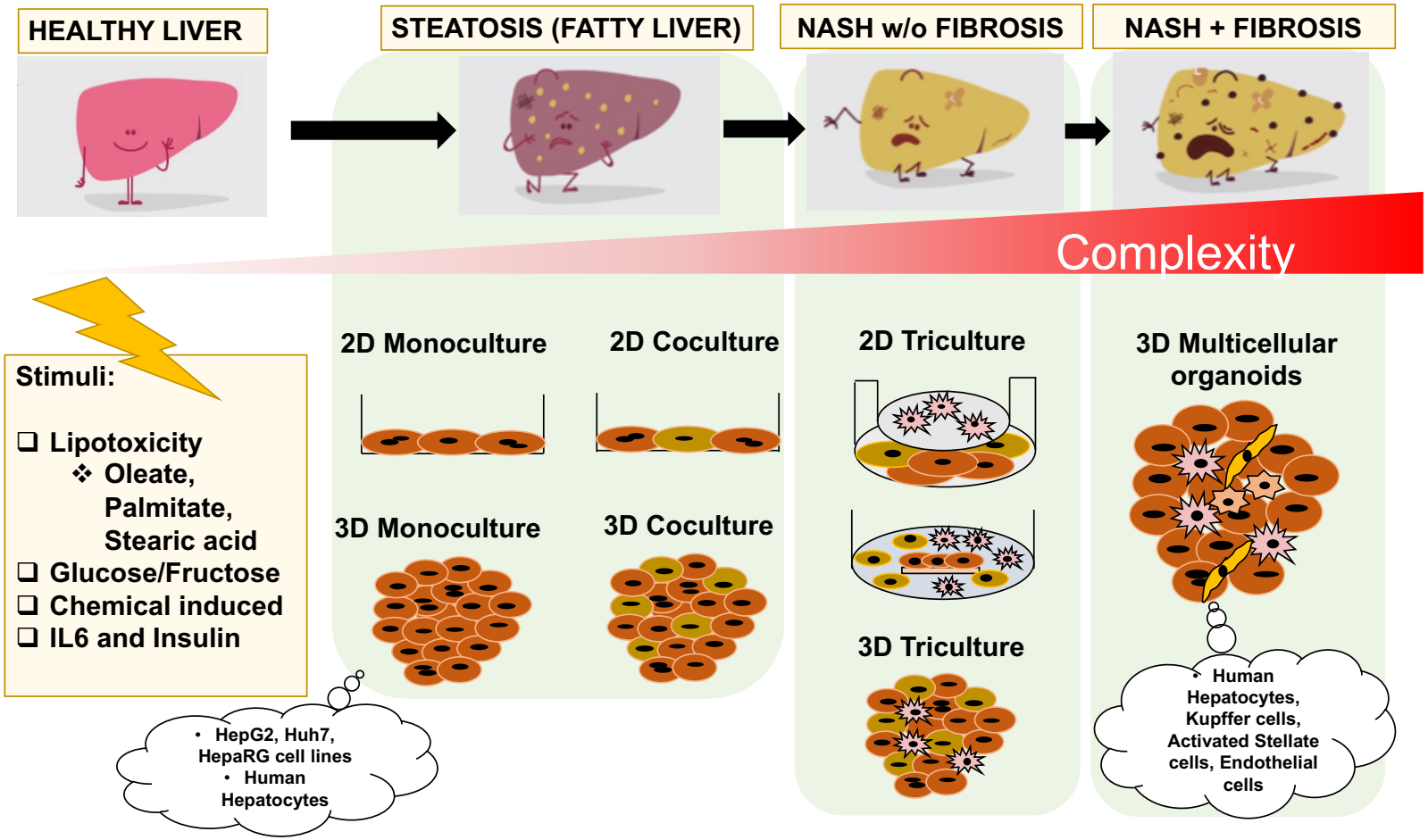

Figure 1. 2: In vitro models of nonalcoholic fatty liver disease.

Abbreviations. NASH: Nonalcoholic steatohepatitis; IL6: Interleukins 6.

Table 1.1: Preclinical in vitro model for NAFLD.

Platform	Treatment	Conc.	Duration	Phenotype	Relevance to human disease	[Ref.]
Monoculture						
HepG2	Oleate	0-1 mM	6 h	↑ Lipid accumulation; ↑ SREBP-1	Hepatic Steatosis	(41)
	Palmitate	0-1 mM	6 h	↑Cytotoxicity; ↑ Insulin resistance; ↑DAG	Progression to NASH	(41)
	Palmitate conjugated to BSA (5:1)	8 mM	24 h	↑ Lipid accumulation; ↑ TG accumulation; ↑IL8 secretion	Progression to NASH	(42)
	Oleate: Palmitate (1:1; 2:1, 0:3) conjugated to BSA	0-2 mM	24 h	↑ Lipid accumulation; ↑TG accumulation	Hepatic Steatosis	(43)
	Valproate	1, 2.5 and 5 µM	24 h	Concentration-dependent increase in intracellular fat; ↑ CD36 and ↑DGAT protein expression.	Hepatic Steatosis	(44)
HepG2 cells culture in microfluidic perfused device	Oleate and Palmitate		24h and 48 h	↓ Lipid accumulation and TG than 2D static model.	NAFLD	(45)
Huh7	Glucose only or Glucose and Fructose (G: F) (1:1)	0.65-0.72 mM or 0.58 - 0.67 mM	24 h	G: F treatment ↑ TG accumulation; ↑ Cholestrol	Hepatic Steatosis	(27)

	Palmitate	0.2 - 0.8 mM	8 h	↑ ER stress	Progression to NASH	(46)
	Oleate: Palmitate (2:1)	0.2 mM	24 h	↑ Lipid accumulation; ↑IL8; ↑IL6 and TNF alpha secretion	NASH	(47)
	Oleate: Palmitate (1:1)	0 - 1.2 mM	24 h	↑ Lipid accumulation	Hepatic Steatosis	(46)
HepaRG	6 negative control and 8 steatotic compounds	Based on the compound	24 h	Steatotic compounds: ↑ lipid accumulation and ↑ ROS production. ↑ gene expression of EHHADH CROT MTTP ANGPTL3 IGFBP1 FOXA1 SREBP1C PPARA PGC1A.	Hepatic Steatosis	(29)
	Tetracycline and amiodarone	0-0.1 mM of tetracycline; 0-0.02 mM of amiodarone	0 and 14 d	↑ Lipid accumulation; ↑TG accumulation; ↑ADFP and ↑PLIN4 genes	Hepatic Steatosis	(7)
	Bisphenol a	0.2 - 2000 nM	21 d	↑ Lipid accumulation; ↑TG accumulation	Hepatic Steatosis	(6)
	Stearic acid or Oleic acid with insulin treatment	0 - 0.1 mM	7 d	↑TG and lipid-responsive genes; Stearic acid ↑ CYP2E1 activity	NAFLD and Obesity	(26)

	Oleate: Palmitate (2:1); Palmitate alone	0.5 mM	48 h	↑ Lipid accumulation; ↑TG; ↑ carnitine metabolism; Palmitate alone contributed ↑ insulin resistance; Palmitate alone induced ↑ oxidative stress;	O:P induced hepatic steatosis; Palmitate induced apoptosis and progression to NASH	(48)
	Oleate: Palmitate (1:1) along with IL6	0.125 mM + 10 ng/mL of IL-6	14 d or 24 h	IL-6 in 24 h ↓ CYP 1A2, 2A6, 2B6, 2C19, 2C9, 3A4 and 3A5 gene expression; O/P in 14 d ↓ CYP 1A2, 2A6, 2C19 gene expression	NAFLD and NASH	(28)
Primary Human Hepatocytes (PHH)	Oleate: Palmitate (2:1)	0.25 - 3 mM	14 h	↓ CYP 1A2, 2A6, 2C9, 2D6, 2E1 and 3A4 mRNA and enzyme activity	Hepatic Steatosis	(32)
3D spheroids from PHH	Oleate: Palmitate (1:1) with varying concentration of insulin, glucose and fructose	160 - 320 µM	48-72h	↑ Lipid accumulation; ↑TG accumulation with steatosis reversibility in absence of FFA; insulin resistant	NASH and NAFLD	(33)
Cryopreserved primary human hepatocytes cultured into 3D model using scaffolds and maintained in LiverChip platform	Oleate: Palmitate (2:1)	0.6 mM	14 d	↓ CYP 3A4 and 2C9 enzyme activity; ↑ Lipid accumulation; ↑TG accumulation	NAFLD/ Steatosis	(34)
Human induced pluripotent stem cell-derived hepatocytes (hiPSC-Hep)	Oleate and Palmitate; including inducer for ER stress	50 - 200 µM (Palmitate); 25 µM Oleate	18 h	↑ Lipid accumulation; ↑TG accumulation	Hepatic Steatosis	(8)

Precision Cut Human Liver Slices	Glucose (G), Fructose (F), insulin (I) and Palmitic acid (PA)	25 mM (G), 5 mM (F), 1 mM (I), 240 μM (PA)	48 h	↑ Lipid accumulation; ↑TG accumulation	Hepatic Steatosis	(40)
Co-culture						
Micro-patterned co-culture (MPCC): Primary Human hepatocytes plated in collagen coated islands along with 3T3-J2 fibroblasts	Glucose	0.5 mM - 25 mM	10 d and 18 d	↑ Albumin secretion; ↑ Urea synthesis; ↑ lipid accumulation; ↑ CYP3A4, 2A6, 2D6 mRNA and ↓ CYP2E1 mRNA; selective insulin resistance.	Hepatic Steatosis and T2DM	(35)
Micro-patterned co-culture (MPCC-HSC): Primary Human hepatocytes plated in collagen coated islands along with hepatic stellate cells (HSCs).	90K HSC:30K PHHs in 24 well format		14 d	↑ CYP3A4 and 2A6 activity; ↑ albumin secretion; ↑ distortion of hepatocyte islands in 2 weeks.	--	(36)
Tri-culture						

Transwell multiculture model: Nonparenchymal cells (NPCs) containing stellate and macrophage (top of transwell) and hepatocytes (bottom of the transwell) incorporated with hepatic sinusoidal flow	Insulin, glucose, Free Fatty Acids.	Insulin (6.9 nM), glucose (25 mM), FFAs (65 μM sodium oleate and 45 μM palmitic acid).	10 d	↑Lipid Accumulation; ↑ *De novo* lipogenesis and cholesterol synthesis; ↑ IL-8 and CK-18; and ↓ insulin signaling	NASH	(37)
Micro-patterned co-culture (MPTC): Primary Human hepatocytes plated in collagen coated islands along with 3T3-J2 fibroblasts and hepatic stellate cells (HSCs) in inserts.	Glucose	5 mM	14 d	↑ Albumin secretion; ↑ Urea synthesis; ↑ lipid accumulation; ↑ CYP3A4, PXR, CAR mRNA and ↓ CYP3A4 and 2A6 enzyme activity; ↓ NTCP, MRP2 and BSEP mRNA; ↑ IL6 and C-RP secretion and selective insulin resistance.	NASH/ fibrosis	(36)
Multi-cell culture						

3D-human liver micro-tissues (InSphero): co-culturing primary human hepatocytes with non-parenchymal cells such as hepatic stellate cells, Kupffer cells and endothelial cells	Palmitate	0.5 mM	9 d	↑ IL8 imaging; ↑ pdgfr***β*** gene expression; ↑ TGFβ pathway; ↑ Collagen deposition	NASH/ fibrosis	(38)
iPSCs derived organoids: 11 healthy and diseased pluripotent stem cells used in developing reprogrammed organoids composed of the hepatocyte, stellate, biliary-like and the kuffer cells.	Oleate	0.8 mM	5 d	↑ Lipid accumulation; ↑TG accumulation; ↑IL6; ↑ hepatocyte ballooning; ↑ inflammation and fibrosis associated genes; ↑ stiffness associated with collagen deposition	Steatosis/ inflammation/ fibrosis	(39)

6. References

1. Bedossa P. Pathology of non-alcoholic fatty liver disease. Liver Int 2017;37 Suppl 1:85-89.

2. Benedict M, Zhang X. Non-alcoholic fatty liver disease: An expanded review. World J Hepatol 2017;9:715-732.

3. Sayiner M, Koenig A, Henry L, Younossi ZM. Epidemiology of Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis in the United States and the Rest of the World. Clin Liver Dis 2016;20:205-214.

4. Filozof C, Goldstein BJ, Williams RN, Sanyal A. Non-Alcoholic Steatohepatitis: Limited Available Treatment Options but Promising Drugs in Development and Recent Progress Towards a Regulatory Approval Pathway. Drugs 2015;75:1373-1392.

5. Hegade VS, Speight RA, Etherington RE, Jones DE. Novel bile acid therapeutics for the treatment of chronic liver diseases. Therap Adv Gastroenterol 2016;9:376-391.

6. Bucher S, Jalili P, Le Guillou D, Begriche K, Rondel K, Martinais S, Zalko D, et al. Bisphenol a induces steatosis in HepaRG cells using a model of perinatal exposure. Environ Toxicol 2017;32:1024-1036.

7. Antherieu S, Rogue A, Fromenty B, Guillouzo A, Robin MA. Induction of vesicular steatosis by amiodarone and tetracycline is associated with up-regulation of lipogenic genes in HepaRG cells. Hepatology 2011;53:1895-1905.

8. Parafati M, Kirby RJ, Khorasanizadeh S, Rastinejad F, Malany S. A nonalcoholic fatty liver disease model in human induced pluripotent stem cell-derived hepatocytes, created by endoplasmic reticulum stress-induced steatosis. Dis Model Mech 2018;11.

9. Jennings P. "The future of in vitro toxicology". Toxicol In Vitro 2015;29:1217-1221.

10. Hartung T, FitzGerald RE, Jennings P, Mirams GR, Peitsch MC, Rostami-Hodjegan A, Shah I, et al. Systems Toxicology: Real World Applications and Opportunities. Chem Res Toxicol 2017;30:870-882.

11. Brunt EM, Janney CG, Di Bisceglie AM, Neuschwander-Tetri BA, Bacon BR. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. Am J Gastroenterol 1999;94:2467-2474.

12. Bedossa P. Histological Assessment of NAFLD. Dig Dis Sci 2016;61:1348-1355.

13. Burt AD, Lackner C, Tiniakos DG. Diagnosis and Assessment of NAFLD: Definitions and Histopathological Classification. Semin Liver Dis 2015;35:207-220.

14. Schwimmer JB, Middleton MS, Behling C, Newton KP, Awai HI, Paiz MN, Lam J, et al. Magnetic resonance imaging and liver histology as biomarkers of hepatic steatosis in children with nonalcoholic fatty liver disease. Hepatology 2015;61:1887-1895.

15. Karlas T, Petroff D, Garnov N, Böhm S, Tenckhoff H, Wittekind C, Wiese M, et al. Non-invasive assessment of hepatic steatosis in patients with NAFLD using controlled attenuation parameter and 1H-MR spectroscopy. PLoS One 2014;9:e91987.

16. Wieckowska A, Zein NN, Yerian LM, Lopez AR, McCullough AJ, Feldstein AE. In vivo assessment of liver cell apoptosis as a novel biomarker of disease severity in nonalcoholic fatty liver disease. Hepatology 2006;44:27-33.

17. Kanwar P, Kowdley KV. The Metabolic Syndrome and Its Influence on Nonalcoholic Steatohepatitis. Clin Liver Dis 2016;20:225-243.

18. Mohanty SR, Troy TN, Huo D, O'Brien BL, Jensen DM, Hart J. Influence of ethnicity on histological differences in non-alcoholic fatty liver disease. J Hepatol 2009;50:797-804.

19. Anstee QM, Day CP. The Genetics of Nonalcoholic Fatty Liver Disease: Spotlight on PNPLA3 and TM6SF2. Semin Liver Dis 2015;35:270-290.

20. Gan L, Chitturi S, Farrell GC. Mechanisms and implications of age-related changes in the liver: nonalcoholic Fatty liver disease in the elderly. Curr Gerontol Geriatr Res 2011;2011:831536.

21. Hamabe A, Uto H, Imamura Y, Kusano K, Mawatari S, Kumagai K, Kure T, et al. Impact of cigarette smoking on onset of nonalcoholic fatty liver disease over a 10-year period. J Gastroenterol 2011;46:769-778.

22. Godos J, Federico A, Dallio M, Scazzina F. Mediterranean diet and nonalcoholic fatty liver disease: molecular mechanisms of protection. Int J Food Sci Nutr 2017;68:18-27.

23. Macut D, Tziomalos K, Božić-Antić I, Bjekić-Macut J, Katsikis I, Papadakis E, Andrić Z, et al. Non-alcoholic fatty liver disease is associated with insulin resistance and lipid accumulation product in women with polycystic ovary syndrome. Hum Reprod 2016;31:1347-1353.

24. Daltro C, Cotrim HP, Alves E, de Freitas LA, Araújo L, Boente L, Leal R, et al. Nonalcoholic fatty liver disease associated with obstructive sleep apnea: just a coincidence? Obes Surg 2010;20:1536-1543.

25. Araya J, Rodrigo R, Videla LA, Thielemann L, Orellana M, Pettinelli P, Poniachik J. Increase in long-chain polyunsaturated fatty acid n - 6/n - 3 ratio in relation to hepatic steatosis in patients with non-alcoholic fatty liver disease. Clin Sci (Lond) 2004;106:635-643.

26. Michaut A, Le Guillou D, Moreau C, Bucher S, McGill MR, Martinais S, Gicquel T, et al. A cellular model to study drug-induced liver injury in nonalcoholic fatty liver disease: Application to acetaminophen. Toxicol Appl Pharmacol 2016;292:40-55.

27. Windemuller F, Xu J, Rabinowitz SS, Hussain MM, Schwarz SM. Lipogenesis in Huh7 cells is promoted by increasing the fructose: Glucose molar ratio. World J Hepatol 2016;8:838-843.

28. Tanner N, Kubik L, Luckert C, Thomas M, Hofmann U, Zanger UM, Böhmert L, et al. Regulation of Drug Metabolism by the Interplay of Inflammatory Signaling, Steatosis, and Xeno-Sensing Receptors in HepaRG Cells. Drug Metab Dispos 2018;46:326-335.

29. Tolosa L, Gómez-Lechón MJ, Jiménez N, Hervás D, Jover R, Donato MT. Advantageous use of HepaRG cells for the screening and mechanistic study of drug-induced steatosis. Toxicol Appl Pharmacol 2016;302:1-9.

30. Gerets HH, Tilmant K, Gerin B, Chanteux H, Depelchin BO, Dhalluin S, Atienzar FA. Characterization of primary human hepatocytes, HepG2 cells, and HepaRG cells at the mRNA level and CYP activity in response to inducers and their predictivity for the detection of human hepatotoxins. Cell Biol Toxicol 2012;28:69-87.

31. Shi J, Wang X, Lyu L, Jiang H, Zhu HJ. Comparison of protein expression between human livers and the hepatic cell lines HepG2, Hep3B, and Huh7 using SWATH and MRM-HR proteomics: Focusing on drug-metabolizing enzymes. Drug Metab Pharmacokinet 2018;33:133-140.

32. Donato MT, Lahoz A, Jiménez N, Pérez G, Serralta A, Mir J, Castell JV, et al. Potential impact of steatosis on cytochrome P450 enzymes of human hepatocytes isolated from fatty liver grafts. Drug Metab Dispos 2006;34:1556-1562.

33. Kozyra M, Johansson I, Nordling Å, Ullah S, Lauschke VM, Ingelman-Sundberg M. Human hepatic 3D spheroids as a model for steatosis and insulin resistance. Sci Rep 2018;8:14297.

34. Kostrzewski T, Cornforth T, Snow SA, Ouro-Gnao L, Rowe C, Large EM, Hughes DJ. Three-dimensional perfused human. World J Gastroenterol 2017;23:204-215.

35. Davidson MD, Ballinger KR, Khetani SR. Long-term exposure to abnormal glucose levels alters drug metabolism pathways and insulin sensitivity in primary human hepatocytes. Sci Rep 2016;6:28178.

36. Davidson MD, Kukla DA, Khetani SR. Microengineered cultures containing human hepatic stellate cells and hepatocytes for drug development. Integr Biol (Camb) 2017;9:662-677.

37. Feaver RE, Cole BK, Lawson MJ, Hoang SA, Marukian S, Blackman BR, Figler RA, et al. Development of an in vitro human liver system for interrogating nonalcoholic steatohepatitis. JCI Insight 2016;1:e90954.

38. Mukherjee S, Zhelnin L, Sanfiz A, Pan J, Li Z, Yarde M, McCarty J, et al. Development and validation of an in vitro 3D model of NASH with severe fibrotic phenotype. Am J Transl Res 2019;11:1531-1540.

39. Ouchi R, Togo S, Kimura M, Shinozawa T, Koido M, Koike H, Thompson W, et al. Modeling Steatohepatitis in Humans with Pluripotent Stem Cell-Derived Organoids. Cell Metab 2019;30:374-384.e376.

40. Palma E, Doornebal EJ, Chokshi S. Precision-cut liver slices: a versatile tool to advance liver research. Hepatol Int 2019;13:51-57.

41. Lee JY, Cho HK, Kwon YH. Palmitate induces insulin resistance without significant intracellular triglyceride accumulation in HepG2 cells. Metabolism 2010;59:927-934.

42. Joshi-Barve S, Barve SS, Amancherla K, Gobejishvili L, Hill D, Cave M, Hote P, et al. Palmitic acid induces production of proinflammatory cytokine interleukin-8 from hepatocytes. Hepatology 2007;46:823-830.

43. Gómez-Lechón MJ, Donato MT, Castell JV, Jover R. Human hepatocytes in primary culture: the choice to investigate drug metabolism in man. Curr Drug Metab 2004;5:443-462.

44. Bai X, Hong W, Cai P, Chen Y, Xu C, Cao D, Yu W, et al. Valproate induced hepatic steatosis by enhanced fatty acid uptake and triglyceride synthesis. Toxicol Appl Pharmacol 2017;324:12-25.

45. Gori M, Simonelli MC, Giannitelli SM, Businaro L, Trombetta M, Rainer A. Investigating Nonalcoholic Fatty Liver Disease in a Liver-on-a-Chip Microfluidic Device. PLoS One 2016;11:e0159729.

46. Takahara I, Akazawa Y, Tabuchi M, Matsuda K, Miyaaki H, Kido Y, Kanda Y, et al. Toyocamycin attenuates free fatty acid-induced hepatic steatosis and apoptosis in cultured hepatocytes and ameliorates nonalcoholic fatty liver disease in mice. PLoS One 2017;12:e0170591.

47. Chavez-Tapia NC, Rosso N, Tiribelli C. Effect of intracellular lipid accumulation in a new model of non-alcoholic fatty liver disease. BMC Gastroenterol 2012;12:20.

48. Brown MV, Compton SA, Milburn MV, Lawton KA, Cheatham B. Metabolomic signatures in lipid-loaded HepaRGs reveal pathways involved in steatotic progression. Obesity (Silver Spring) 2013;21:E561-570.

In Preparation for submission to Molecular & Cellular Proteomics

MANUSCRIPT 2

COMPARATIVE PROTEOMIC ANALYSIS OF CRYOPRESERVED PRIMARY HEPATOCYTES AND HEPATIC CELL LINES USING SWATH-MS REVEAL SIGNIFICANT VARIATIONS IN PROTEINS INVOLVED IN ENERGY, LIPID AND XENOBIOTIC METABOLISM

Anitha Saravanakumar[1], Benjamin B Barlock[1], Rohitash Jamwal[1], Xin Bush[1], Scott Heyward[2], Fatemeh Akhlaghi[1]

1 Biomedical and Pharmaceutical Sciences, College of Pharmacy, University of Rhode Island, Kingston, RI, 02881, USA

2 BioreclamationIVT, 1450 South Rolling Road Baltimore, MD 21227, USA

Corresponding author at:

Fatemeh Akhlaghi, Ph.D.
Clinical Pharmacokinetics Research Laboratory
Department of Biomedical and Pharmaceutical Sciences
The University of Rhode Island, 495A College of Pharmacy, 7 Greenhouse Road,
Kingston, RI 02881, United States.
Email address: fatemeh@uri.edu

Abbreviations

DMEs - Drug metabolizing enzymes

DDA - Data-dependent acquisition

TOF - Time of flight

CYP450 - Cytochrome P450

UGT - UDP-glucuronosyltransferase

SULT - Sulfotransferase

ACADSB - Acyl-CoA dehydrogenase short/branched chain

ACSM3 - Acyl-CoA synthetase medium-chain family member 3

ACSM5- Acyl-CoA synthetase medium-chain family member 5

CPS1 - Carbamoyl-phosphate synthase

MS -Mass spectrometry

LFQ - Label-free quantification

DTT - 1,4-Dithiothreitol

IAA- Iodoacetamide

DOC - Sodium deoxycholate

Conflict of interest

None of the authors have any conflict of interest to declare.

Statement of financial support

Financial support for this study was provided by National Institutes of Health grants to Fatemeh Akhlaghi [grant numbers R15-GM101599, UH3-TR000963].

Abstract

Background & Aim. Human hepatic carcinoma cell lines are commonly used in invitro studies of lipid and xenobiotic metabolism, as well as glucose regulation in normal and disease state. However, their validity is still under debate due to the variable expression of proteins in the cell lines and human hepatocytes. In this present study, we used a data independent acquisition based total protein approach (DIA-TPA) to quantify the protein abundance in the different cell lines versus (vs.) human hepatocytes (cHH) and human liver tissue (HLT). For this purpose, the global proteome from the whole cell homogenates of HepaRG, HepG2, Huh7 cell lines were compared to cHH and HLT.

Methods. In-solution trypsin digestion of proteins in whole cell homogenate was conducted using pressure-cycling technology (PCT). DIA was carried out by sequential window acquisition of theoretical mass spectra (SWATH-MS). In addition, MS2 spectra for all detectable peptides at its corresponding *m/z* were quantified using Spectronaut™, which was then analyzed using TPA.

Results. In summary, 2715, 2578, 2874, 2717 and 3083 proteins in, HepaRG, HepG2, Huh7, cHH and HLT, were identified at 1% FDR, respectively. The global proteome of the cHH significantly differed from the cancer hepatic cell lines. Within the cell lines, the global as well as ADME protein profile of HepaRG most closely correlated with cHH, with 84 out of 101 ADME proteins, identified in HepaRG cells. Within gluconeogenesis and glycolysis pathway, Huh7 cell line expressed proteins in high abundance in contrast to the other groups.

Conclusion. In summary, the comparison demonstrates the capability of untargeted global proteomics to detect the differences in protein expression among the different groups. In addition, this study provides a comprehensive database of information to aid better study design.

Keywords.

HepaRG, Huh7, HepG2, Human Hepatocytes, Drug metabolizing enzymes, CYP450, Proteomics, SWATH-MS.

1. Introduction

The human liver is the major organ for metabolism and clearance of xenobiotics, maintenance of glucose homeostasis as well as regulation of lipid metabolism.[1] Majority of the circulating proteins in the plasma, such as albumin, that constitutes 55-70% of human plasma is produced by the liver. Liver also play a central role in the excretion of bile and cholesterol. The diverse metabolic and secretory function of the liver is aided by the presence of specialized cell types such as hepatocytes, kupffer, stellate and biliary endothelial cells. However, 80% of the tissue is comprised of the hepatocytes alone.[2] The hepatocytes are mostly binucleated[3] and contains abundant proteins localized to endoplasmic reticulum and mitochondria, indicating their effective role in protein synthesis and energy metabolism, respectively.

In hepatocytes, the glucose consumption is observed via the uptake of plasma glucose, followed by the break-down of the sugar moiety to pyruvate, which in-turn is oxidized through the TCA cycle to produce ATP. Conversely, the glucose may also get converted to fatty acids via de-novo lipogenesis pathways.[4] On the other hand, hepatocytes also play a critical role in lipid metabolism via conversion of non-esterified fatty acids into tri-acylglycerols (TAG), which in-turn is stored or secreted as low-density lipoproteins (VLDL) back into the plasma.[4] Studies characterizing the mechanisms involved in metabolic functioning of the liver, in the healthy as well as disease state such as obesity, insulin resistance, non-alcoholic fatty liver disease (NAFLD) are on the increase.[5-8] The gold standard for such invitro studies are human hepatocytes. However, the cost of procurement, donor-to-donor variability as well as limited availability renders some disadvantages of using them for long-term studies. Consequently, hepatic carcinoma cell lines such as HepG2, Huh7 and HepaRG cells are often used as an alternative tool. In addition, immortalized hepatic cell lines have been routinely been used

in toxicity as well as drug metabolism assays in preclinical studies.[9-11] However, it is unclear how well the tumor-derived cell lines represent the human hepatocytes in-context to their protein expression involved in lipid and energy metabolism as well as xenobiotic metabolism.

Shotgun proteomics is a powerful tool for analyzing proteolytic peptides of high intensity, from low-throughput biological assays. The results generated are high-throughput and have a complete proteome coverage that out-weighs the traditional labor and resource intensive western blotting techniques.[2,12] Data-dependent acquisition (DDA) has been used more commonly in untargeted analysis.[13] However, bias associated with the abundant peptides and the lack of sensitivity to quantify low-abundance proteins, renders limitation to use this approach.[14] Hence, sequential window acquisition of all theoretical mass spectra (SWATH-MS) that follows a Data Independent Acquisition (DIA) was employed to analyze the samples. Eventually, the DIA based total protein approach (DIA-TPA) was used to compare the absolute protein levels of the human hepatocytes and the carcinoma cell lines.[15-17] This approach is sensitive to capture the low abundant proteins which makes it ideal to compare the expression pattern of thousands of proteins from the cell lines and human hepatocytes, at the same time. In addition, we compared the protein expression from whole cell homogenate without enrichment. We believe this novel approach of sample analysis overcomes the limitation of loss of proteins in fractionation and enrichment and is more reflective of the biological state.[2,18,19]

The objective of this study was to compare the protein expression involved in the glucose, lipid as drug metabolism pathways in immortalized hepatocyte cell-lines as well as primary human hepatocytes. This comprehensive comparison provides a database to aid the betterment of study design in the field of hepatology and pathophysiology.

2. Materials and Methods

2.1. Chemical and Reagents

Trypsin digested β-galactosidase (E. coli) (Catalog No. #4445247) and TPCK-treated trypsin (Catalog No. #4445250) was procured from Sciex (Framingham, MA). Bovine serum albumin (BSA) (Catalog No. #23209) and Iodoacetamide (IAA) (Catalog No. #I6125), Dithiothreitol (DTT) (Catalog No. #DTT-RO), Ammonium bicarbonate (Catalog No. #40867) and Sodium deoxycholate (DOC) (Catalog No. #30970) were purchased from Sigma-Aldrich (St. Louis, MO). MS grade acetonitrile (Catalog No. #A955), methanol (Catalog No. #A456), choloroform (Catalog No. #C297) and formic acid (Catalog No. #A117) were from Fisher-Scientific (Hampton, NH). Acquity UHPLC Peptide BEH C18 analytical column (Catalog No. #186005594) and VanGuard precolumn (Catalog No. #186004629) were from Waters Corp. (Milford, MA).

2.2. Cell Culture

The cell lines and cHH were cultured as per conditions in-house or as per manufacturer's protocol. The conditions between the different cell lines and cHH were maintained as close as possible with minor changes that conformed to specific cell lines and cHH.

Cryopreserved primary human hepatocytes (cHH). Three donors of cHH, namely lot, LQG, AQL, VHB were obtained from Bio ReclamationIVT (Baltimore, MD) (Table 2.1) were thawed in InVitro GRO CP media (Bio ReclamationIVT, Catalog No. #Z99029, Baltimore, MD) and viable cells were counted using Tryphan-Blue (Sigma-Aldrich, Catalog No. #T8154, St. Louis, MO) exclusion method. The cells were seeded in collagen coated 6-well plate (Corning Inc., Catalog No. #354400, Corning, NY) at a seeding density of 0.7 million cells per well. After 4 h of plating, cells were washed and incubated using CP media, to remove cell debris. Following 24 h of incubation at 37°C, cells were extracted by gentle scraping using cell scraper (CELLTREAT Scientific Products, Catalog No.

#229310, Pepperell, MA) in ice-cold DPBS (ThermoFisher Scientific, Catalog No. #14040117, Waltham, MA) and 10-million cells from each lot was pooled for further analysis. It was found that the cell attachment efficiency varied from donor to donor. However, all lots were approximately 90% confluent before cell extraction.

HepG2 cells. HepG2 cells were purchased from American Type Culture Collection (Manassas, VA). The HepG2 cells were cultured in Eagle's Minimum Essential Media (Sigma-Aldrich, Catalog No. #M2279, St. Louis, MO) supplemented with 10% FBS (Sigma-Aldrich, Catalog No. #F2442, St. Louis, MO), 1% penicillin/streptomycin (ThermoFisher Scientific, Catalog No. #15140122, Waltham, MA). Cell passage (6 & 7) was performed with Trypsin-EDTA (ThermoFisher Scientific, Catalog No. #25200072, Waltham, MA) when HepG2 cells were 80% confluent in culture. The cells were cultured in 25 cm^2 flask (Sigma-Aldrich, Catalog No. #CLS430639, St. Louis, MO) for 24h, before subjected to cell extraction for proteomics analysis. Approximately 8-10 million cells were cultured and harvested by cell scraping.

HuH7 cells. Huh7 cell line was kindly provided by Dr. Nisanne Ghonem (University of Rhode Island). Cells were cultured in Dulbecco's Modified Eagle Media (ThermoFisher Scientific, Catalog No. #11320082, Waltham, MA) supplemented with FBS and 1% penicillin/streptomycin. The concentration of FBS was maintained at 10% until cells reached confluence. Upon confluence, cells were maintained for 3 weeks in 1% FBS.[11] Cell passages (4 & 5) were performed with 0.25% trypsin-EDTA when cells reached 90% confluence in 25 cm^2 flask. Approximately 8-10 million cells were cultured and harvested by cell scraping.

HepaRG cells. Frozen undifferentiated HepaRG cells were purchased from Biopredic International (Rennes, France). The cells were grown for 2 weeks in William's E media (ThermoFisher Scientific, Catalog No. #12551032, Waltham, MA), containing 1%

glutamine (ThermoFisher Scientific, Catalog No. # A1286001, Waltham, MA), 1% penicillin/streptomycin and 10% HepaRG growth supplement (Lonza Inc., Catalog No. #ADD711C, Walkersville, MD) in 6-well plate. Following growth phase, cells were differentiated in-house for 2 weeks in William's E media, 1% glutamine, 1% penicillin/streptomycin and 10% HepaRG differentiation supplement (Lonza Inc., Catalog No. #ADD721C, Walkersville, MD) and maintained at 37°C with 21% O_2 and 5% CO_2 throughout the culture with media changed every 3 days. Upon differentiation, HepaRG cells were cultured in basal media containing William's E media, 100 U/mL penicillin, 100 µg/mL streptomycin, 2 mM glutamine, 5 µg/mL insulin (Sigma-Aldrich, Catalog No. #I-034, St. Louis, MO) and 50 µM hydrocortisone hemisuccinate (Sigma-Aldrich, Catalog No. #H2270, St. Louis, MO). Approximately 8-10 million cells were cultured and harvested by cell scraping for proteomic analysis.

Human Liver Tissue. 100 mg of frozen human liver samples (n=3) from brain dead donors were purchased from Sekisui XenoTech LLC (Kansas City, KS). The detailed demographics of the donors are given in Table 2.1. Livers were graded by a histopathologist as previously described[20] and were categorized as normal, healthy livers.

2.3. Whole cell homogenate and proteomics sample preparation

An overview of the study workflow is given in Fig. 1. Cells scraped and collected in DPBS was centrifuged at 300g at 4°C for 10 min. The cell pellets obtained was washed in DPBS as mentioned above. Approximately 8-10 million cells from each group was pooled and homogenized in 1000 µl of homogenization buffer (8 M urea, 50 mM Triethylammonium bicarbonate, 10 mM DTT v/v) using a bead homogenizer (Omni Bead Ruptor 24, Kennesaw, GA). In addition, 50 mg of liver tissue was homogenized in 1000 µl of homogenization buffer, in similar fashion. Samples were then spun at 1,000 *g* for 5 min and the supernatant was collected. The total protein concentration of the resulting sample

was determined using Pierce BCA protein assay kit (ThermoFisher Scientific, Catalog No. #23227, Waltham, MA).

2.4. In-solution trypsin digestion

Protein digestion was conducted as described previously with few adaptations.[21,22] Samples (250 µg protein) were spiked with 2 µg of BSA and denatured with 25 µL of DTT (100 mM) at 35°C for 30 min in a shaking water bath (100 rpm) (Precision Scientific, Catalog No. #66800, Chicago, IL). After denaturation, samples were alkylated in the dark with 25 µL of IAA (200 mM) for 30 min at room temperature. Samples were then concentrated using the cold water, methanol and chloroform (1:2:1) precipitation method (centrifugation at 10000 rpm, 5 min at 10°C) (Eppendorf, Catalog No. #5810 R, Hamburg, Germany). The protein pellet was washed with ice-cold methanol and suspended in 100 µL of 50 mM ammonium bicarbonate (pH ~8) containing 3% w/v sodium deoxycholate (DOC). Further, TPCK-treated trypsin (10 µg) was added to samples at a ratio of 1:25 (trypsin: protein) and samples were transferred into digestion tubes (PCT MicroTubes, Pressure Biosciences Inc., Easton, MA). The barocycler was run at 35°C, for 75 cycles with 60-sec pressure-cycle (50-sec high pressure, 10-sec ambient pressure, 25 kpsi). Subsequently, 10 µg trypsin was again added to each sample and digestion was repeated as mentioned above.

Furthermore, to 110 µL of digested peptides sample, 10 µL of acetonitrile (1:1, v/v containing 5% formic acid) was added to precipitate DOC. Samples were spun (10,000 rpm for 5 min at 10°C) to remove the precipitate and 100 µL supernatant was collected. Subsequently, 25 µL of the digested peptide sample was injected on the analytical column and were analyzed using LC-MS/MS method described below.

2.5. LC-MS/MS Analysis

Data-dependent analysis (DIA) was performed in positive ionization mode using a DuoSpray™ ion source on a Sciex 5600 TripleTOF™ mass spectrometer (AB Sciex, Concord, Canada) equipped with an Acquity UPLC HClass system (Waters Corp., Milford, MA, USA). Gas 1 (GS1), gas 2 (GS2) and curtain gas (CUR) were maintained at 55, 60 and 25 psi, respectively. Ion spray voltage floating (ISVF) was kept at 5500 V while the source temperature (TEM) was 500°C. Declustering potential (DP), collision energy (CE) and collision energy spread (CES) were set at 120, 10 and 5 respectively. During the survey scan, all the ions with a charge state of 2 to 4, mass range of *m/z* 300-1250 and exceeding 25 cps were used for MS/MS analysis. Former target ions were excluded for 8 sec and the mass tolerance for TOF-MS was 50 mDa with a 100 milliseconds accumulation time. For product scan, data were acquired from 100 to 1250 *m/z* with an accumulation time of 75 milliseconds with a total cycle time of 3.5 sec. Product ion analysis was done under dynamic accumulation and rolling collision energy dependent on the *m/z* of the ion.

Chromatographic separation was achieved over 180 min gradient method at 100 µL/min on an Acquity UPLC Peptide BEH C18 (2.1 X 150 mm, 300 Å, 1.7 µm) preceded by an Acquity VanGuard pre-column (2.1 X 5 mm, 300 Å, 1.7 µm). Mobile phase A was 98% water, 2% acetonitrile, 0.1% formic acid and mobile phase B was 98% acetonitrile, 2% water, 0.1% formic acid). Gradient conditions used were 98% A from 0 to 5 min, 98% to 70% A from 5 to 155 min, 70% to 50% A from 155 to 160 min, 50% to 5% A from 160 to 170 min, 5% to 98% A held from 170 to 175 min. The gradient was held at initial conditions from 175 min until the end of the run to equilibrate the column before the start of next run. The flow was diverted to waste for the first 8 minutes and last 20 minutes of the acquisition. Autosampler was maintained at 10°C, and the column was kept at

50°C. Trypsin-digested β-galactosidase peptides were injected to monitor TOF detector mass calibration every four samples.

2.6. SWATH-MS acquisition and data analysis

Homogenate samples were analyzed using SWATH-MS based spectra which were acquired for mass range m/z 400−1100 Da within SWATH window width of 10 m/z resulting in 70 overlapping mass windows per cycle, as described in the previous literature[21]. The DIA data was searched against Swiss-Prot identifiers (October 2016) and analyzed using Spectronaut™ Pulsar software (version 11.0, Biognosys AG, Schlieren, Switzerland) was used to obtain MS2 signal intensities of fragment ions from DIA data using default settings with a few modifications that includes specific trypsin/P digestion with a minimum of 7 amino acids, maximum of 52 amino acids and 2 missed cleavages; fixed modifications included carbamidomethyl (C) and variable modifications included acetyl (protein N-terminus) and oxidation (M). Raw data files and search results are available at Japan Proteome Standard Repository (jPOSTrepo JPST000372, ProteomeXchange PXD008593). The SWATH-MS DIA data was matched with the reference spectral library generated from the DDA of pooled human liver samples using Protein Pilot. The DIA data was then normalized and analyzed using "Total protein approach" (TPA), as shown below (1).

$$\text{Protein(i)} = \frac{MS\ signal(i)}{Total\ MS\ signal \times \text{Molecular mass(i)}} \times 10^9 \quad \ldots (1)$$

The absolute protein abundance of each group was represented as pmol/mg of protein.

2.7. Statistical and bioinformatic analysis

A number of normalization strategies were tested using NormalyserDE. LIMMA package with multiple comparison was used to test for significance (FDR < 0.05, log2FC > 0.58 or

log2FC <0.58) to identify differentially expressed proteins. Differentially expressed pathways associated between human hepatocytes, and the three hepatocarcinoma cell lines was assessed using ingenuity pathway analysis (Ingenuity Systems, Redwood City, CA, USA). PCA analysis was conducted using Perseus (version 14.1, MKS Umetrics). Hierarchical clustering was calculated using Euclidean distance metric was conducted using Spectronaut™ Pulsar software. Venn diagram was drawn using an online tool (http://bioinformatics.psb.ugent.be/webtools/Venn/). GraphPad Prism (version 8.0) was used in data visualization and statistics. The Phase I, II and transporters protein expression among the different groups were analyzed using one-way ANOVA and post-hoc analysis of multiple comparison Bonferroni test was carried out to assess significance ($P<0.05$).

3. Results

3.1. Global proteome profile comparisons from DIA data among the whole cell lysate of human hepatocytes, human liver tissue as well as the HepaRG, HepG2 and Huh7 cell lines

As illustrated in the overall workflow (Fig. 2.1), DIA data was first generated from the whole cell lysate samples of human hepatocytes (lot LQG, AQL, VHB), human liver tissue (510, 758 and 791) and three hepatic cell lines, namely, the HepaRG, HepG2 and Huh7. As a result, 2715, 2578, 2874, 2717 and 3083 proteins in, HepaRG, HepG2, Huh7, cHH and HLT, were identified at 1% FDR, respectively (Table 2.2). It is to noted that, 153 (7.2%) proteins in HepaRG was found to be common with HLT and not with cHH (Fig. 2.2B), and we believe this is due to the bipotent progenitor status of HepaRG, resulting in expression of hepatocyte-like and biliary-like cells.

In addition, the global proteome profiles of the cell lines, cHH and HLT were analyzed using Perseus and represented as a PCA and heat map (Fig. 2.3B). Dramatic differences were observed between cHH, HLT and the three cell lines. With regard to the PCA analysis of component 1 and 2, it was noted that the HLT and cHH were grouped closer to one another in comparison to the hepatic carcinoma cell lines. In addition, hierarchical clustering (Fig. 2A) showed a similar pattern of the HLT and cHH grouped together with minimum distance and hepatic carcinoma cell lines were clustered closer to one another.

3.2. SWATH-MS analysis of ADME protein expression in the HLT, cHH and the hepatic cell lines

To investigate the detection reliability, we performed label-free quantification on both HLT and cHH (Fig. 2.4A). More than 2000 proteins were identified, of which 97%

(2647 out of 2717) were found in both sample types. Overall, a good global correlation in protein abundance of individual proteins were observed between the two sample sets. In addition, we detected 89 out of the 101 ADME proteins (88%) chosen from the literature. The ADME protein abundance, showed a good correlation around the line of unity.

The hepatic cell lines correlated moderately with the hepatocytes with 2309 (85%), 2174 (80%), 2308 (84%) of the global proteins from HepaRG, HepG2 and Huh7 matched with 2717 proteins of cHH, respectively. Similarly, we detected 92 (91%), 70 (69%) and 74 (73%) out of the 101 ADME proteins, respectively. HepaRG, showed a more comparable levels of ADME proteins (Fig. 2.4B), however Huh7 and HepG2 showed a distribution of ADME proteins below the line of unity showing lower expression in these cell lines in comparison to hepatocytes (Fig. 2.4C and D).

3.3. SWATH-MS analysis of proteins involved in energy and lipid metabolism in the HLT, cHH and hepatic cell lines

Hepatocytes are pivotal in the energy supply and storage. The major pathways such as glycolysis/gluconeogenesis, are involved in the conversion glucose to pyruvate, eventually for ATP production (Fig. 2.5). Glucose undergoes phosphorylation by hexokinase or glucokinase to form glucose-6-phosphate (G6P) that is converted to ATP via glycolysis. Furthermore, via gluconeogenesis, pyruvate is converted to acetyl-CoA that participates in ATP production as seen in TCA cycle. Alternatively, glucose is also used up in lipid synthesis via pentose-phosphate pathway.

The protein expression in terms of energy metabolism varied significantly in HLT, cHH and the hepatic cell lines (Fig. 2.6). For instance, the primary enzyme involved in gluconeogenesis, 1,6-glucose bisphosphatase (FBP1/2), was identified only in cHH and HLT. Other mitochondrial proteins involved in gluconeogenesis such as PCK2 and PC

were found in higher abundance (~6 fold) in cHH than HepaRG cell lines. Hexokinase (HK1), important enzymes involved in glycolysis was present in HLT, cHH and was comparable to the levels in HepaRG cell line. In contrast, LDHA involved in the conversion of lactate (released by the muscles) to glucose in hepatocytes was found in high abundance in Huh7 cell lines (9-fold) in comparison to cHH (Fig. 2.6). There is some evidence in literature, showing the ability of cancer cells to be reprogrammed to utilize glycolysis at higher rates, in support to this we observed higher abundance of majority of enzymes (ENO1, GAPDH, PGAM1, PGK1, PKLR, PKM, TPI1) in Huh7 cell line in comparison to cHH (Table 2.3).[23]

The non-esterified fatty acids (NEFA) released from the adipocytes are taken up by the hepatocytes and ligated to CoA that initiates the metabolism of fatty acids or triacyl glyceride (TAG) synthesis. This process is carried out by different isoforms of ACSL, namely 1,3 and 5. In cHH and HLT the most abundant form was noted to be ACSL1 and the levels were comparable to HepaRG cell line. However, the isoform ACSL3 and ACSL5 play a predominant role in HepG2 cell line (Fig. 2.5). The fatty acids with more than 16 carbon chain may preferentially undergo peroxisomal oxidation. Enoyl-CoA Hydratase and 3-Hydroxyacyl CoA Dehydrogenase (EHHADH), key enzyme involved in fatty acid beta-oxidation was measured as 17.1 pmol/mg which was similar to HLT and 2-fold higher than HepaRG (8.5 pmol/mg) cell line (Table. 2.4) Among the hepatic cell lines, HepaRG cell line showed highest protein abundance of the groups of protein involved in both peroxisomal and mitochondrial beta-oxidation in comparison to Huh7 and HepG2 cells and was more comparable to cHH.

3.4. SWATH-MS analysis of proteins involved in xenobiotic transport and drug metabolism

Liver is the major organ that contributes to clearance and metabolism of drugs and xenobiotics present in the circulation. This is accomplished by the specialized drug

metabolizing enzymes present in the hepatocytes. The drugs in the circulations are taken up by the hepatocytes via basolateral uptake transporters (Fig. 2.9A), followed by phase I and/or phase II metabolism (Fig. 2.7 and 2.8), and eventually the metabolites are excreted back into the blood or bile (Fig. 2.9B) to be eliminated by the kidney or the intestine, respectively. In general, solute carrier (SLC) family facilitate uptake of xenobiotics from the blood, while the efflux of compounds from hepatocytes are carried out by the ATP-binding cassette (ABC) family. It is to be noted, that our findings showed the most abundant uptake transporters detected in cHH and HLT were organic anion transporter 2 (OAT2), organic anion transporting polypeptide 1B1 (OATP1B1), and OATP1B3 at around 4.0, 1.3 and 1.2 pmol/mg of total protein (Fig. 2.9A). OCT 7 as well as OATP1B1 and 2B1 were detected at within quantifiable limits in the HepaRG but not in HepG2 and Huh7. Among the efflux transporters MRP6 was the most abundant followed by MRP2, BSEP and MDR1.

The disposition of phase I and II metabolizing enzymes in cHH, HLT and the hepatic cell lines is outlined in the Fig.2.7 and 2.8. Human hepatocytes expressed all clinically relevant CYP450, such as, CYP1A2, 2B6, 2C9, 2C19, 2D6 and 3A4 at 1-24 pmol/mg of total protein on an average, detected from the three lots of cHH. HepaRG cell line, expressed some of these enzymes in comparable levels, namely, CYP2C19, 2D6, 3A4 and 3A5. In addition, CYP1A2, 2B6, 2C8 and 2C9 were detected at lower levels in HepaRG than cHH. In comparison, HepG2 cell line did not express CYP1A2, 2B6, 2C8, 2C9, 2C19, 3A5 and 4A11 in detectable levels. In exception, to other CYPs, HepG2 cell line was a high expresser for CYP2C19 (32.6 pmol/mg of whole cell lysate). Likewise, it was interesting to note that Huh7 expressed CYP4A11, enzyme involved in fatty acid metabolism, at levels higher than human hepatocytes. Amongst phase II metabolizing enzymes, SULT1A1, 2A1 as well as UGT1A1, 1A3, 1A9, 2B15 were expressed at almost

similar levels in the cHH as seen in HepaRG cells. However, majority of the phase II enzymes detected in HepG2 and Huh7, were low abundance proteins (~1-2 pmol/mg of total protein) in comparison to cHH.

3.5. Differentially regulated biological pathways in cell lines in comparison to human hepatocytes.

IPA predicted the top 5 dysregulated pathways from the protein expression in each group (HepaRG, HepG2, Huh7) vs. human hepatocytes with p value < 0.05 and log FC>0.58 or <-0.58 (Fig. 2.10). It was to be noted that the biological pathways that was perturbed was similar across the groups (Table 3.5). In addition, we also noted that the predicted top upstream regulator that was significantly inhibited across all cell lines, was found to be, HNF4α. We believe this probably contributes to the dysregulated protein expression pattern seen in lipid and xenobiotic metabolism.

4. Discussion

Human hepatic carcinoma cell lines are commonly used in drug development to aid safety assessment and for candidate selection for first-in-human (FIH) studies. Hence, there is a need for well-characterized, fit-to-purpose, proliferative hepatic cell lines models, due to the limited availability of primary human hepatocytes cells, inconsistency across donors and variability in DMET expression.[24] In addition, the need to understand hepatic-disease state along with its impact on drug metabolizing enzymes propels the urgency for better preclinical models of human origin. With this objective in mind, we screened HepaRG cells as well as more traditional cell lines such as HepG2 and Huh7 for the expression of DME and proteins involved in energy and lipid metabolism.[6,7] HepaRG are cells derived from a 66 y old female patient with hepatocellular carcinoma.[26] They are progenitor cells that upon differentiation exhibit 50% hepatocyte-like and/or 50% biliary-like cell lineages.[27] In addition, upon differentiation, HepaRG cells are reported to express high levels of CYP450 and other Phase I and Phase II enzymes, that is confirmed using several micro array studies.[9,28,29] Apart from the reported gene expression data, very few reports discuss the global proteome of these cell lines as well as the targeted DME protein expression. Also, it is important to note that the known gene expression reports may not directly extrapolate to its protein expression or its activity.[30] Proteins being the functional unit of most biological processes, it is critical to understand the protein expression profiles before choosing a hepatic cell line to model drug metabolism or disease state.

To date, only a few proteomic studies have been reported in the literature comparing the global proteomes of hepatic cell lines and human hepatocytes as well as to quantify the DME protein expression. Some of the studies include contributions from: Slany et al. wherein the proteome profiles of primary human hepatocytes were compared to HepG2 and Hep3B using 2D-PAGE approach coupled with DDA based shortgun

proteomics[31]; Xiaomei Gu and colleagues compared the proteomic characteristics of HepG2, Hep3B2, H226, Ovcar3 and N87 using a TPA from Maxquant with DDA approach[13]; Sison-young et al. compared HepG2, Upcyte and HepaRG to primary human hepatocytes using a DDA strategy[32]; and recently Shi and colleagues used a DIA/SWATH based approach to compare HepG2, Hep3B, and Huh7 to human hepatocytes.[14] In most cases, DDA based approach was used in analysis of proteomics signature. However, it has limitations for the quantification of low abundant proteins. In contrast, DIA based SWATH-MS approach is more sensitive and demonstrates high coverage, reproducibility and precision for low abundant proteins. Hence, DIA based SWATH-MS approach is ideal for measuring proteins in cell lines with low-to-moderate protein yield. Using the SWATH-MS detection method we compared the global proteome expression, along with targeted proteins involved in energy, lipid and xenobiotic metabolism among the cell lines and cHH as well as HLT. Moreover, in contrast, to all the above-mentioned literature that employed sub-cellular fractionation for enrichment, we analyzed the whole cell homogenate without fractionation, to prevent the loss of proteins associated with fractionation, as well as the variations associated with the scale-up of the fractionates. Henceforth, we believe, this analysis will result in higher reproducibility and limit the batch-to-batch variability associated with fractionation.

In the current study, a spectral library was generated from the DDA scans. A total of <6000 proteins (combined) was identified using Protein Pilot at 1% FDR. However, not all proteins were detected in all replicates. DIA based total protein approach (DIA-TPA) was used to compare the global proteome among the different groups. We validated this approach by comparing the values of DMET proteins reported in literature as pmol/mg of total protein in wholes cell lysate. Wiśniewski et. al (2016) reported the protein levels from human hepatocytes and HepG2 cell line using DDA based TPA approach. Our data

closely matched with the published literature, despite the different stratergy (DIA-TPA) involved in the analysis (Fig. 2.11-2.13).

Similar to the previous reports, dramatically different global proteome profiles were observed between the cell lines and human hepatocytes (Figure 2.3, 2.4).[14,32] However, hepatic carcinoma cell lines most closely related to each other in comparison cHH and HLT. Among the proteins expressed, the differentially regulated proteins as well as the perturbed pathways among the different groups were analyzed using IPA (Qiagen, Valencia, CA). As reported in the previous literature, the EIF2 signaling pathway, sirutin and MTOR pathway along with biological oxidation were significantly dysregulated in hepatic cell lines as opposed to human hepatocytes.[14,32]

Several DMEs were quantified in the cHH and HLT as well as the cell lines using DIA based TPA approach. Among the 12 CYP450s quantified in cHH and HLT; 11 CYPs were identified in HepaRG; 7 CYPs in Huh7 and 5 CYPs in HepG2 cell line. The abundance of CYPs in HepaRG was more comparable to cHH than HepG2 and Huh7. UGTs were highly expressed in cHH and HLT. In HepaRG cell line, 10 UGTs and 2 SULTs were detected, however, UGTs (1A1, 1A3, 1A4 and 2B15) were not quantifiable in HepG2 and HuH7. Hence caution must be taken while choosing these cell lines for drug metabolism related experiments. Having said that, it is also important to note that HepG2 cells expressed high levels of, MGST2 (2-fold), POR (2.5-fold), DHRS2 (360-fold) in comparison to cHH. These enzymes have shown involvement in xenobiotic metabolism as well as cancer. In contrast, Huh7 cells expressed GSTP1 (75-fold), AKR1C1 (0.7-fold), ALDH1A1 (3-fold) and ADH5 (1.7 fold) in high abundance in comparison cHH (Appendix I). Therefore, these cell lines can be a legitimate in vitro model when experiments are tailored for targeted proteins of interest. [33,34]

With respect to the proteins involved in energy and lipid metabolism, we compared a proteins involved in glycolysis, ß-oxidation and lipogenesis across different groups. Majority of proteins involved in gluconeogenesis/glycolysis was expressed in high abundance in Huh7 cell line (Figure 2.4B). Interestingly, the average expression across the proteins involved in fatty acid metabolism was much higher in the HLT and cHH in comparison to cell lines. HepaRG, expressed a general lower trend than cHH followed by HepG2 and Huh7 cells. However, in contrast to the DME expression, the proteins involved in lipid metabolism was highly expressed in the carcinoma cell lines, but not of the similar abundance as human hepatocytes.

In conclusion, in this publication we used a DIA based TPA strategy to compare proteome profiles of the whole cell homogenate of three hepatocarcinoma cell lines, i.e. HepG2, Huh7 and HepaRG and compared it with cHH and HLT in the focus of proteins associated to pathways of energy, lipid and xenobiotic metabolism. As discussed before, we observed significantly different protein expression profiles among the groups in energy and lipid metabolism as well as ADME protein expression. Therefore, caution must be exercised when choosing the cell lines as in vitro models for drug metabolism or developing models for disease states. This study provides a comprehensive database for understanding the expression of proteins in different cell lines and we believe it will aid in making informed choices for hepatic cell lines in future model development.

5. Figures and Tables

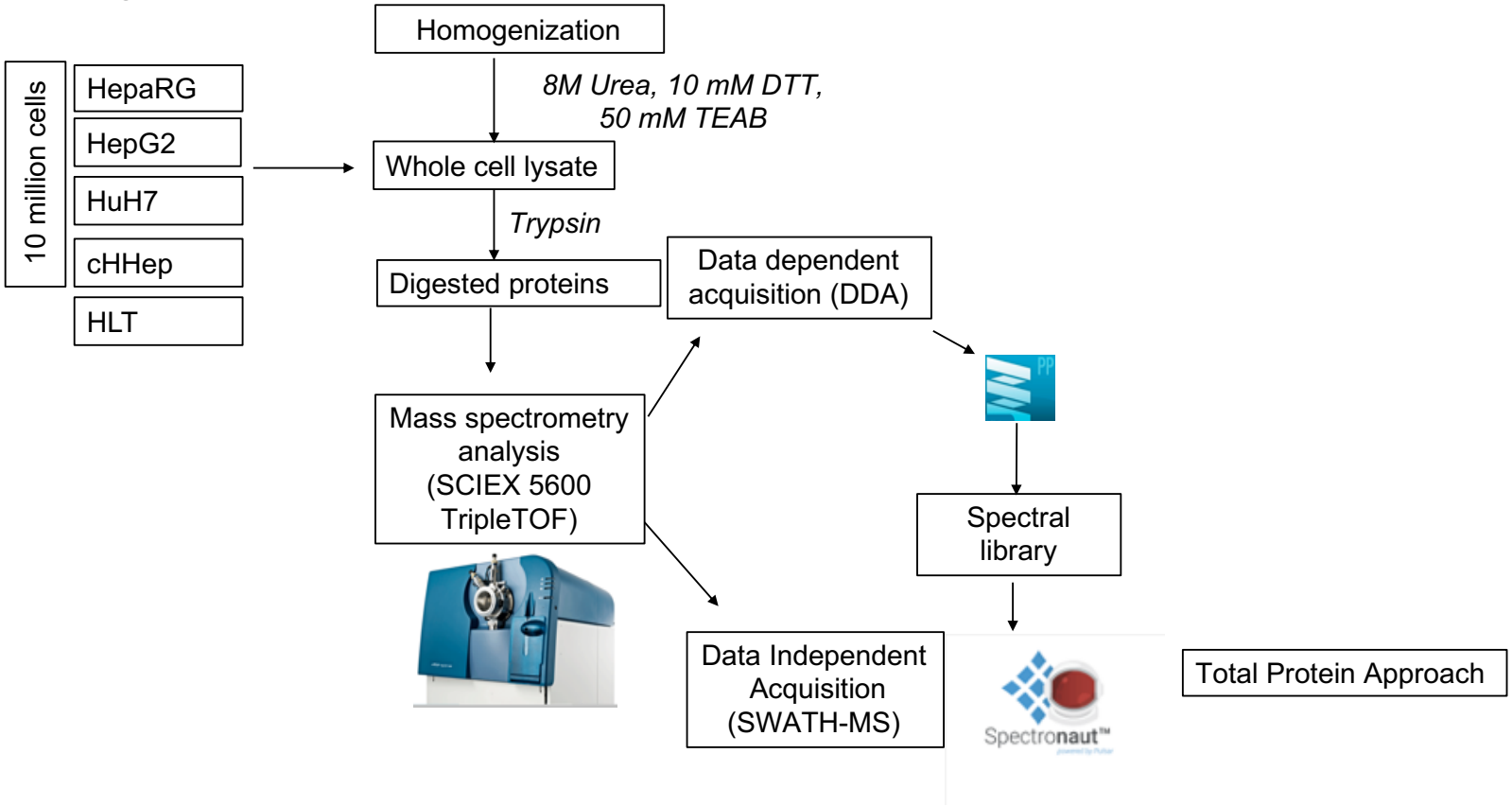

Figure 2.1: The figure illustrates the overall workflow of protein detection, identification and quantitation.

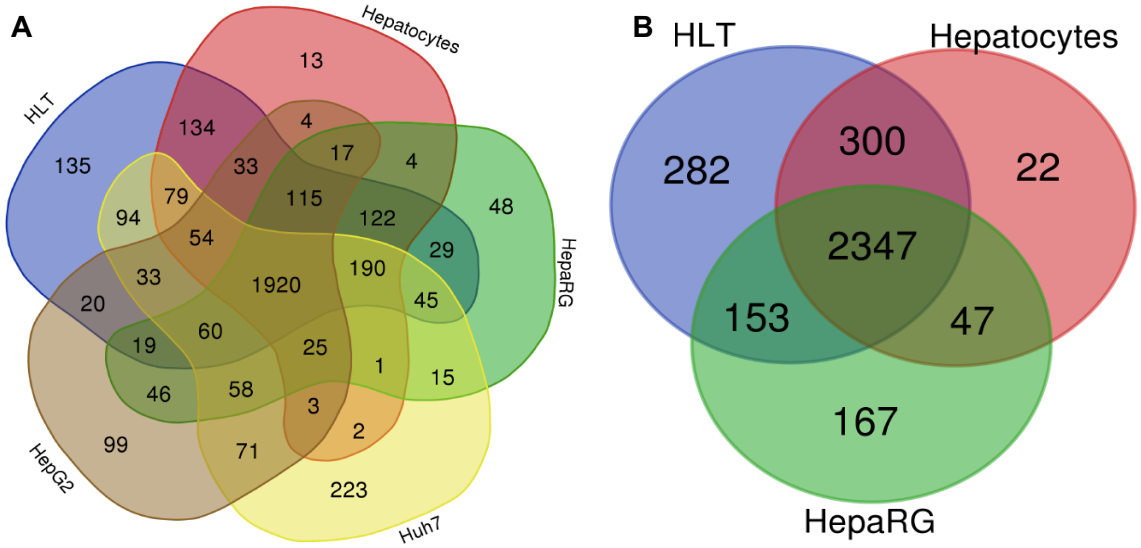

Figure 2.2: Venn diagram representing the overlapping number of proteins.

The figure shows the identified number of proteins in (A) human livers tissue, human hepatocytes and the HepaRG, HepG2, and Huh7 cell lines and (B) HLT, Hepatocytes and HepaRG alone, quantified by SWATH-MS. All proteins were identified in all three replicates for each group. The diagram was produced using a web tool available at http://bioinformatics.psb.ugent.be/webtools/C/. Abbreviation: HLT- Human liver Tissue.

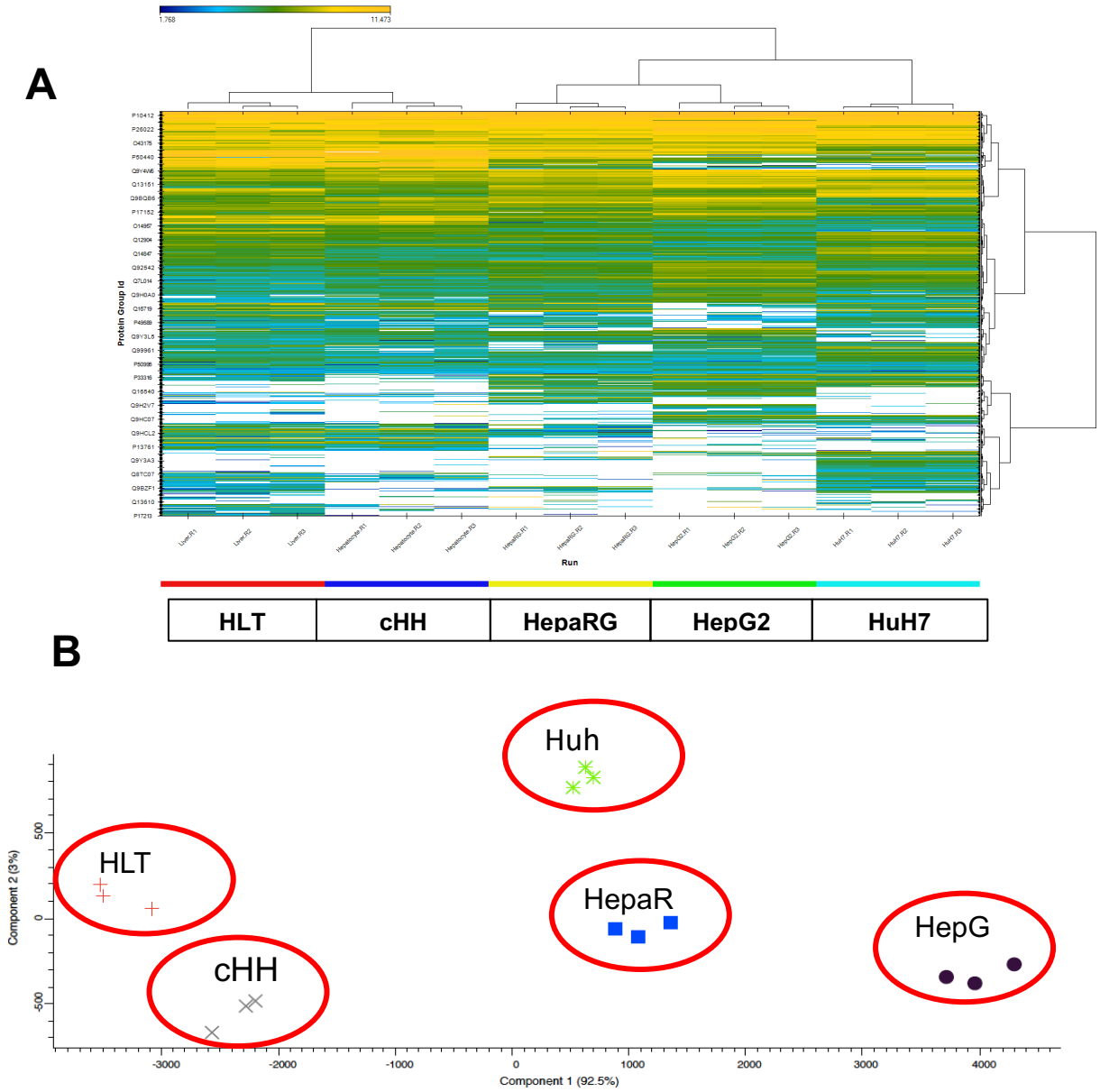

Figure 2.3: Heirarchial clustering of proteins differentially expression.

The figure (A) shows HLT, cHH, and the HepaRG, HepG2, and Huh7 cell lines (n= 3 for each group) clustered together depending on Euclidean distance matrix. The 30 most abundant proteins determined by DIA-TPA are represented in the heat map The MS2 signal intensity was presented as a Log2 value. White spots indicate zero intensity. The figure (B) shows the grouping of cHLT, cHH and the HepaRG, HepG2, and Huh7 cell lines generated from SWATH-MS based data.

Abbreviation: HLT – Human Liver Tissue; cHH – cryopreserved Human Hepatocytes; SWATH-MS - Sequential Window Acquisition of theoretical Mass Spectra; DIA-TPA – Data Independent Acquisition based Total Protein Approach.

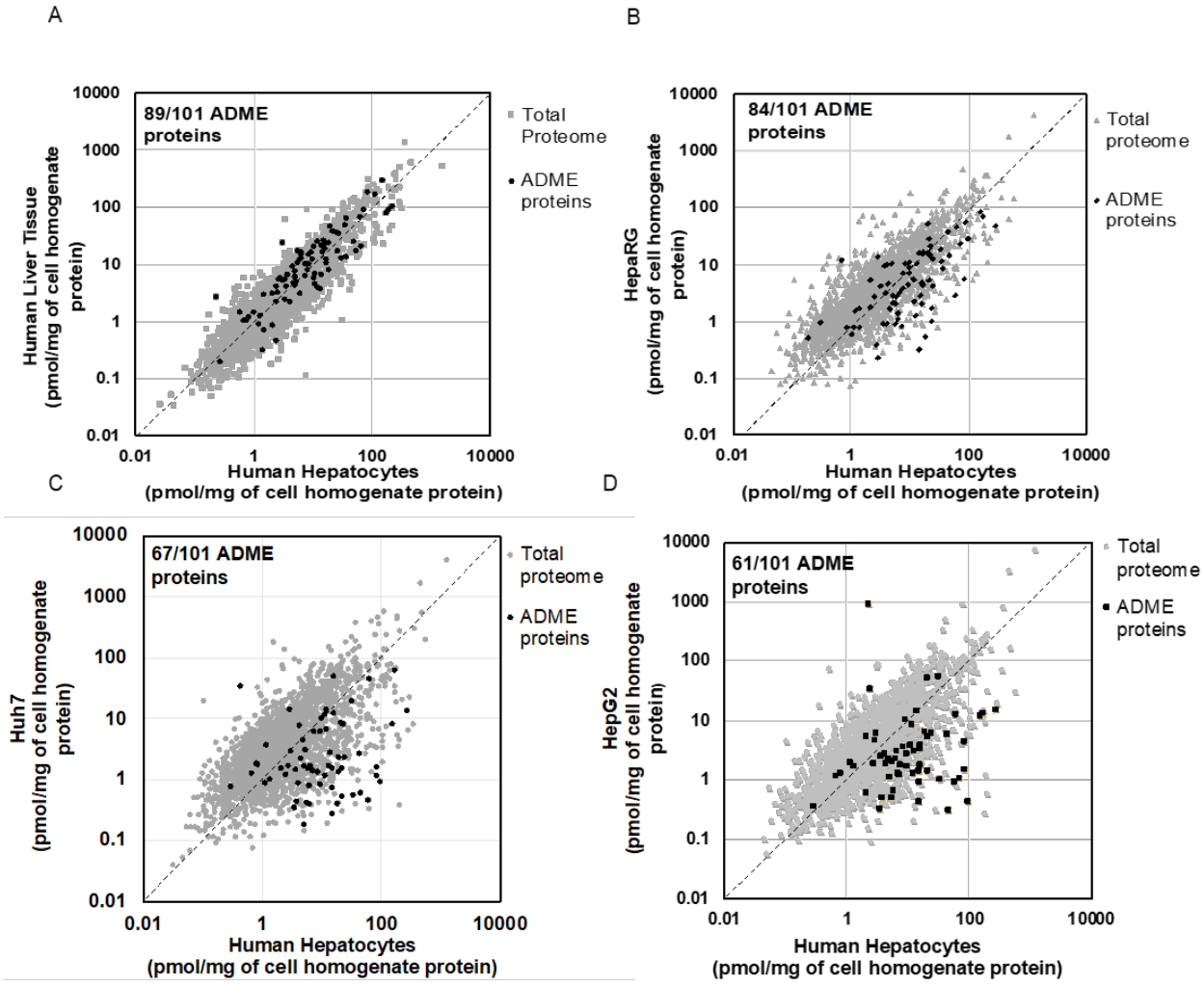

Figure 2.4: Correlation of protein abundances between the two groups.

(A) Human Hepatocytes vs. HLT (B) Human Hepatocytes vs. HepaRG (C) Human Hepatocytes vs. HuH7 (D) Human Hepatocytes vs. HepG2. The global correlation in protein expression indicated in grey and the correlation of ADME proteins in black. Dotted lines indicate the line of unity.

Abbreviation: HLT – Human Liver Tissue.

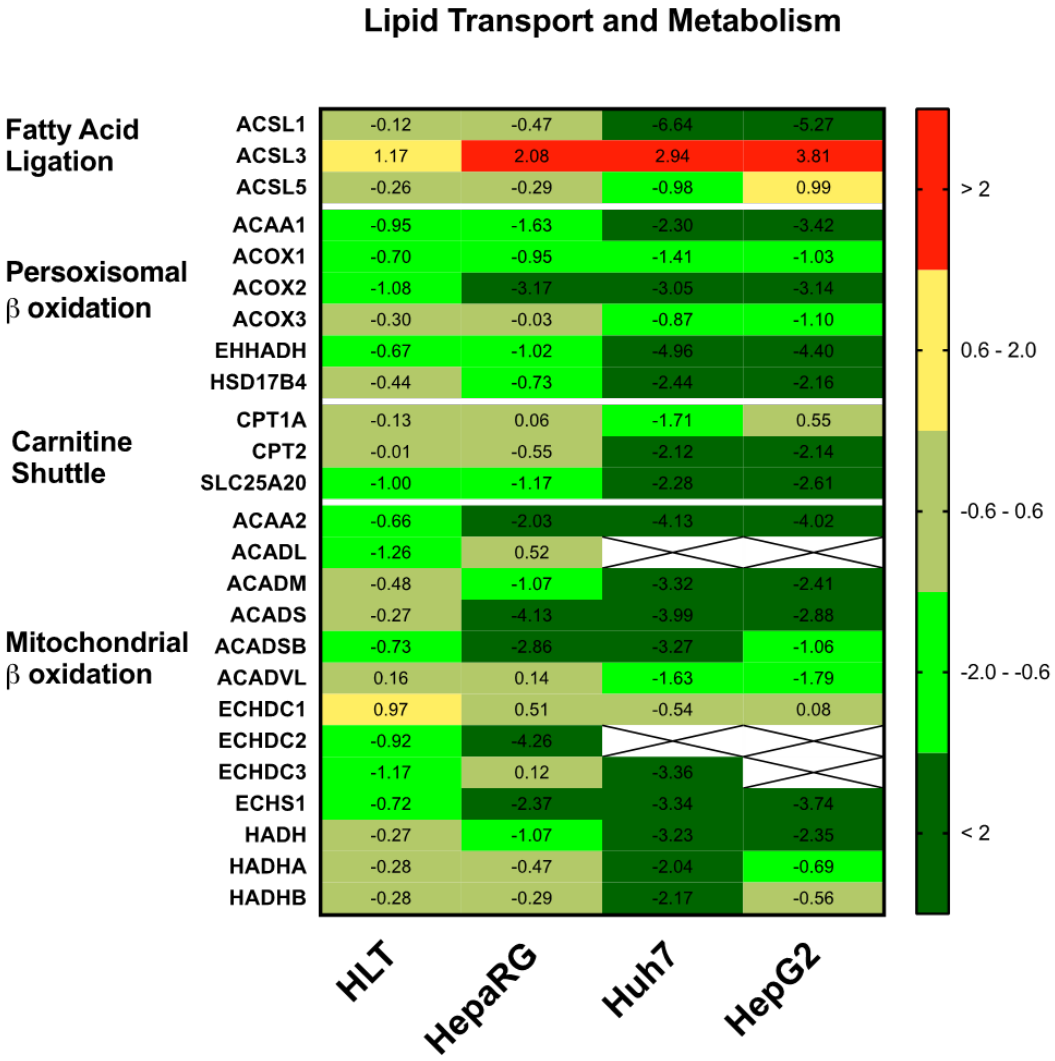

Figure 2.5: Heat map of the differential expression of protein involved in glucose metabolism using SWATH-MS.

The protein expression is represented as the log2 fold change between the test group (HLT, HepaRG, Huh7 and HepG2) vs. human hepatocytes. The figure shows proteins involved in gluconeogenesis/glycolysis. X – not applicable when protein expression is below limit of quantification. Abbreviation: HLT –Human liver tissue.

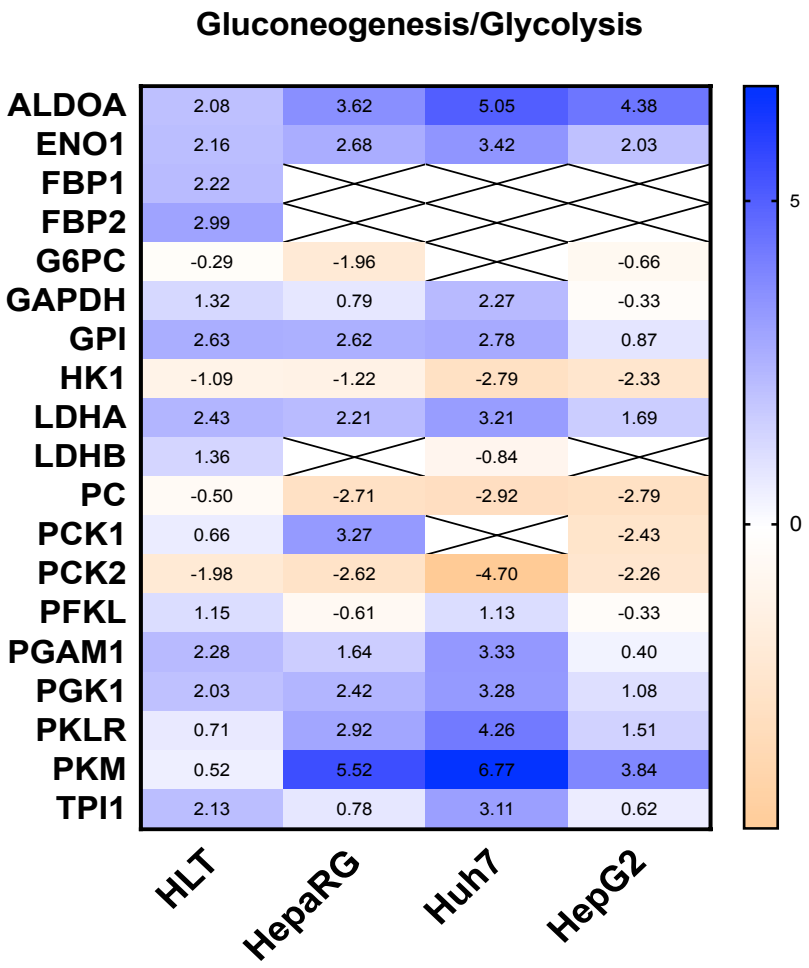

Figure 2.6: Heat map of the differential expression of protein involved in lipid metabolism using SWATH-MS.

The protein expression is represented as the log2 fold change between the test group (HLT, HepaRG, Huh7 and HepG2) vs. human hepatocytes. The figure shows proteins involved in lipid metabolism. X – not applicable when protein expression is below limit of quantification. Abbreviation: HLT –Human liver tissue; FA – Fatty acid.

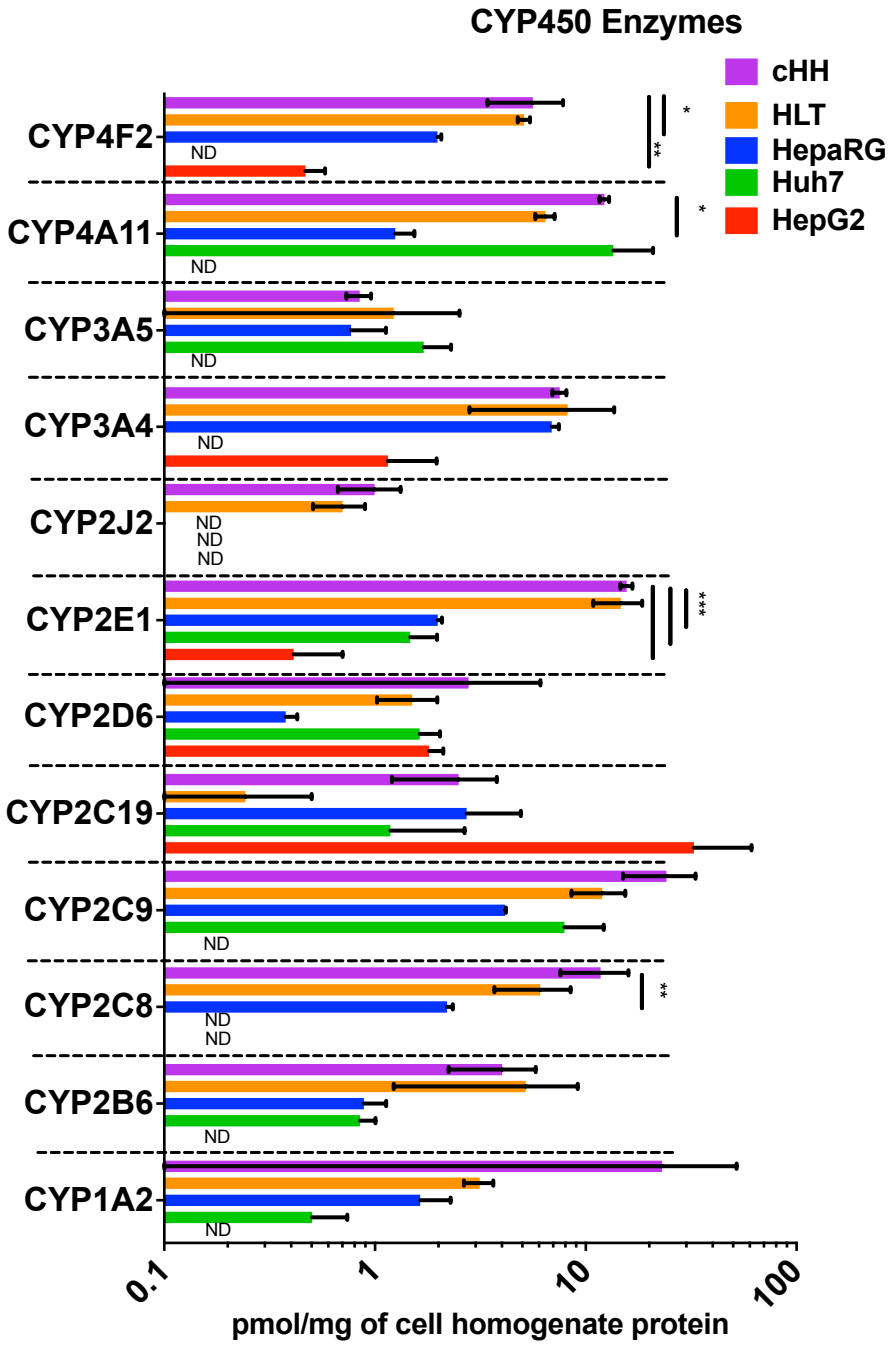

Figure 2.7: Cytochrome P450 expressing profile from SWATH-MS.

The significance was assessed using one-way ANOVA with multiple comparison Bonferroni test * denotes p <0.05; ** denotes p <0.01; *** p<0.001. ND stands for not determined due to low expression. The proteins were expressed as the average of triplicates in pmol/mg of total protein. Error bars represent SD. Abbreviations: CYP - cytochrome.

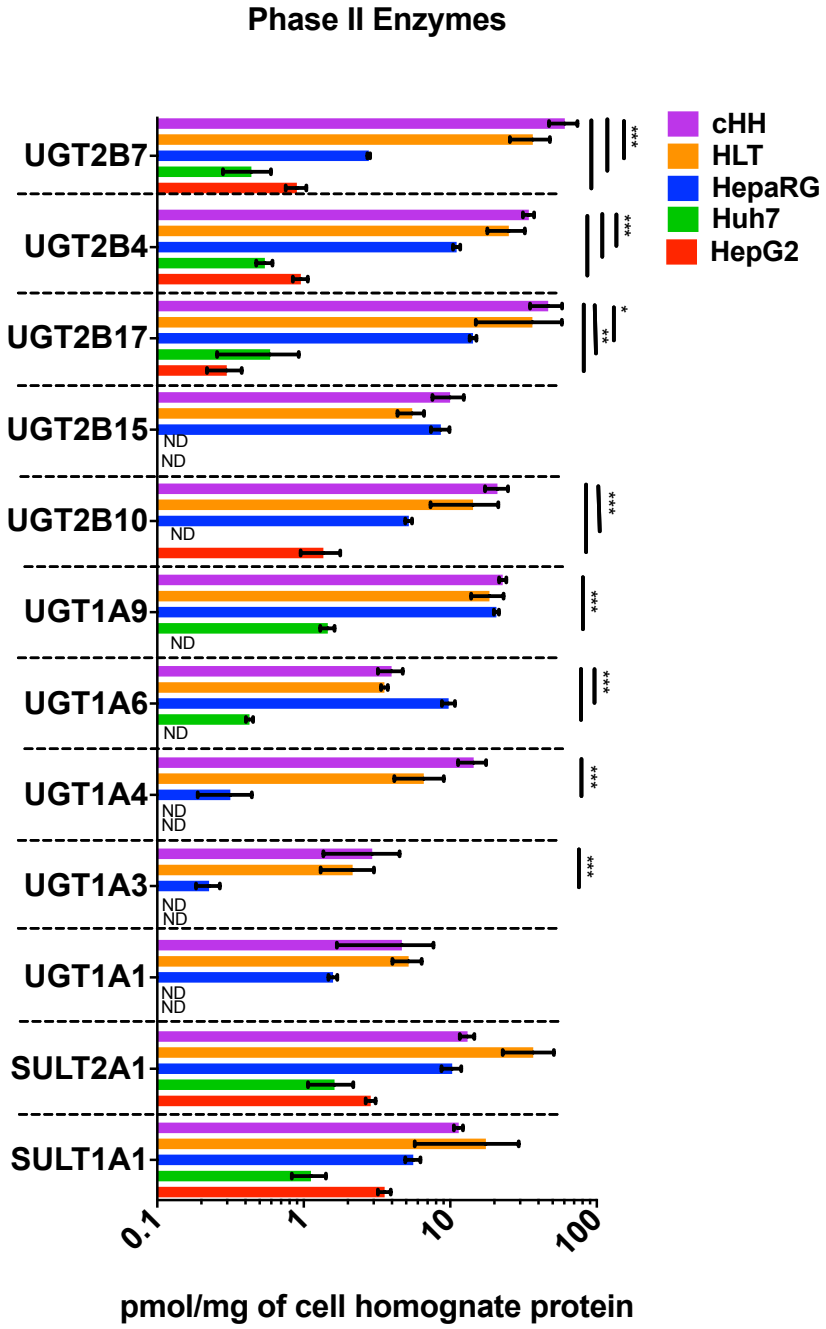

Figure 2.8: Drug metabolizing enzymes involved in phase II.

The significance was assessed using one-way ANOVA with multiple comparison Bonferroni test * denotes p <0.05; ** denotes p <0.01; *** p<0.001. ND stands for not determined due to low expression. The proteins were expressed as the average of triplicates in pmol/mg of total protein. Error bars represent SD. Abbreviations: UGT - UDP-glucuronosyltransferase; SULT – sulfotransferase.

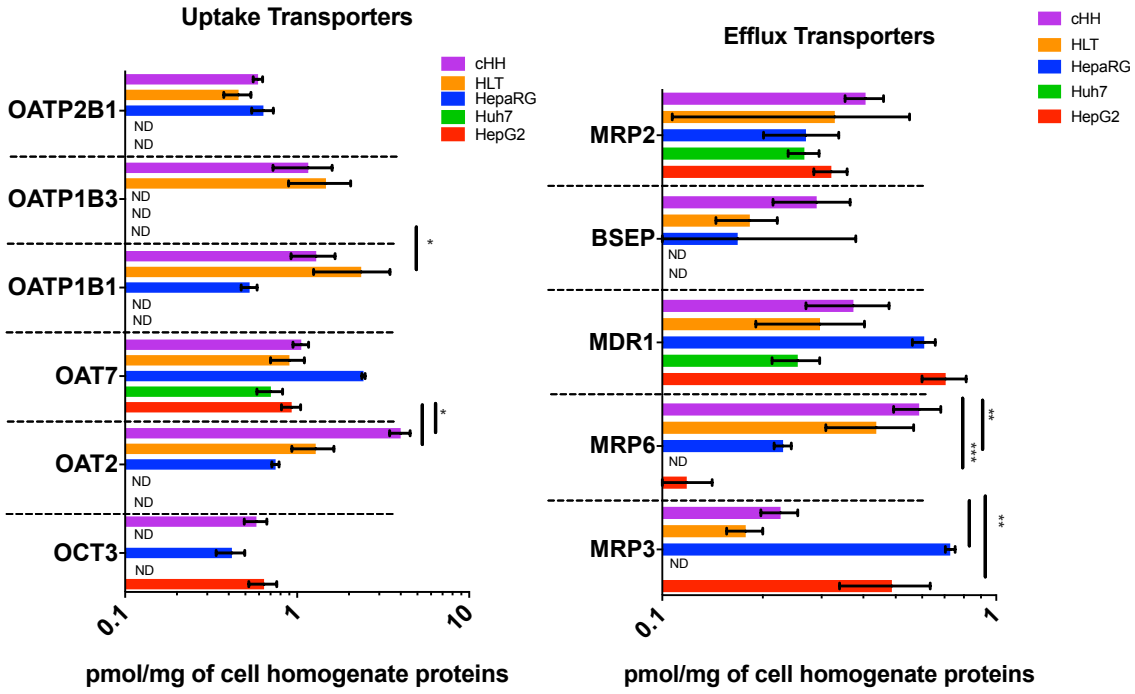

Figure 2.9: Xenobiotic transporters quantified using SWATH-MS.

The significance was assessed using one-way ANOVA with multiple comparison Bonferroni test * denotes p <0.05; ** denotes p <0.01; *** p<0.001. ND stands for not determined due to low expression. The proteins were expressed as the average of triplicates in pmol/mg of total protein. (A) uptake transporters (B) efflux transporters. Error bars represent SD.

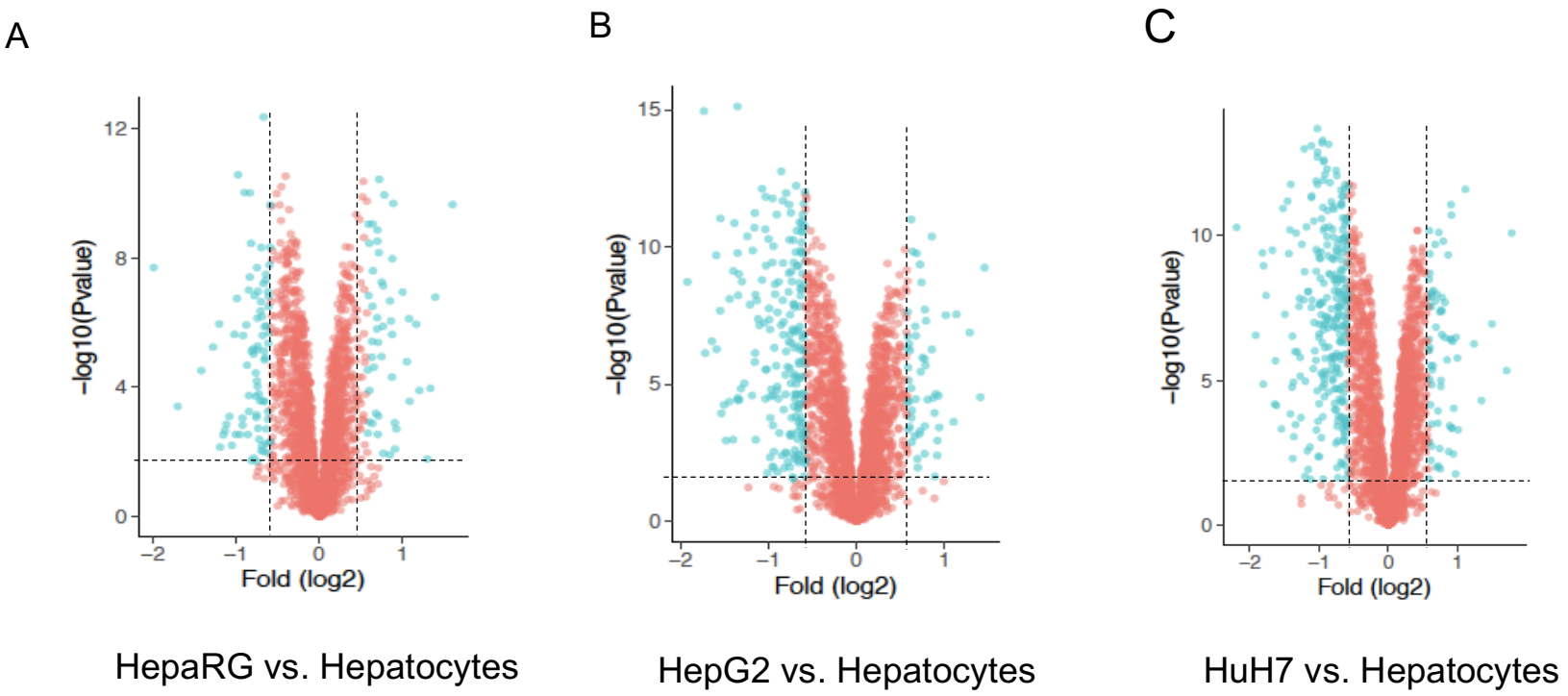

Figure 2. 10: Volcano plots representing number of differentially expressed proteins in HepaRG, HepG2 and Huh7 in comparison to Human Hepatocytes.

Dots in blue represent the differentially expressed proteins based on FDR < 0.05 and log2FC <0.58 or log2FC >0.58 in (A) HepaRG (B) HepG2 (C) HepaRG in comparison to human hepatocytes.

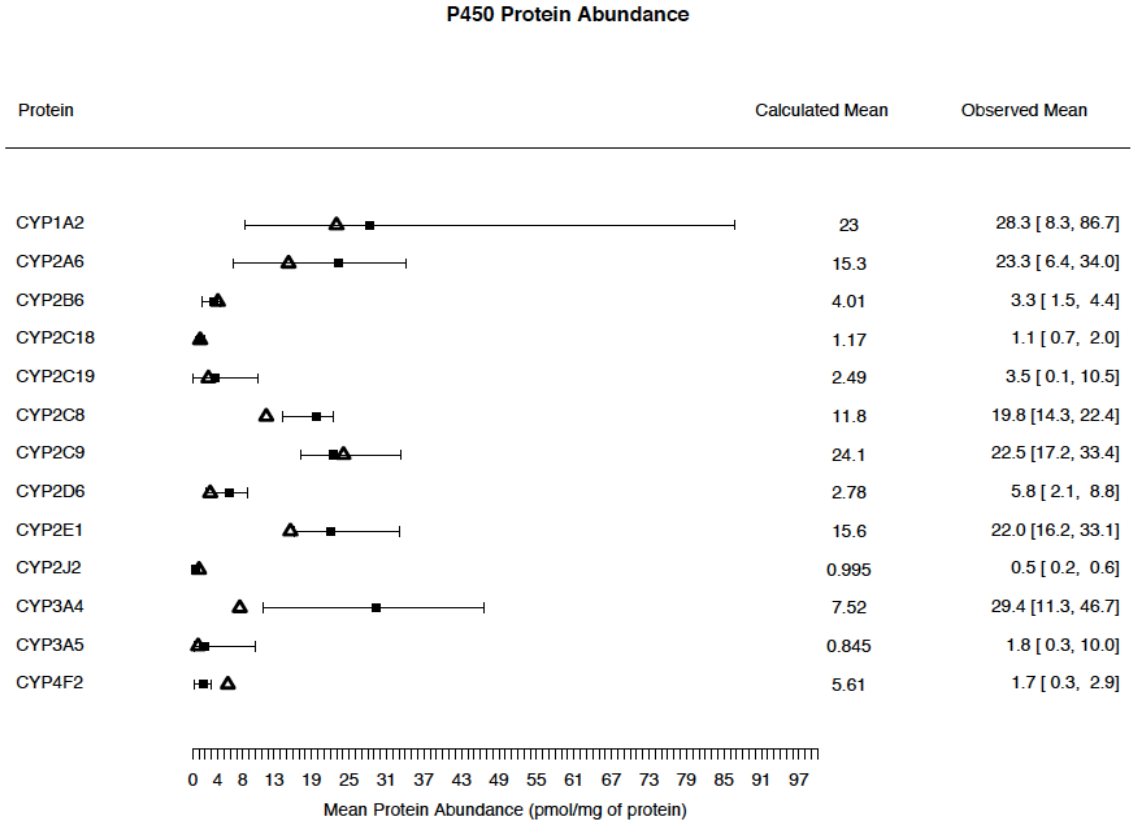

Figure 2.11: Comparison of P450 protein using SWATH-MS DIA-TPA data with published data (Weisnewski et. al, 2016) based on DDA-TPA approach.

The square in black represent the average of human hepatocytes (n=6) enzymes with min and max range. The triangle represents the average of human hepatocytes from current study.

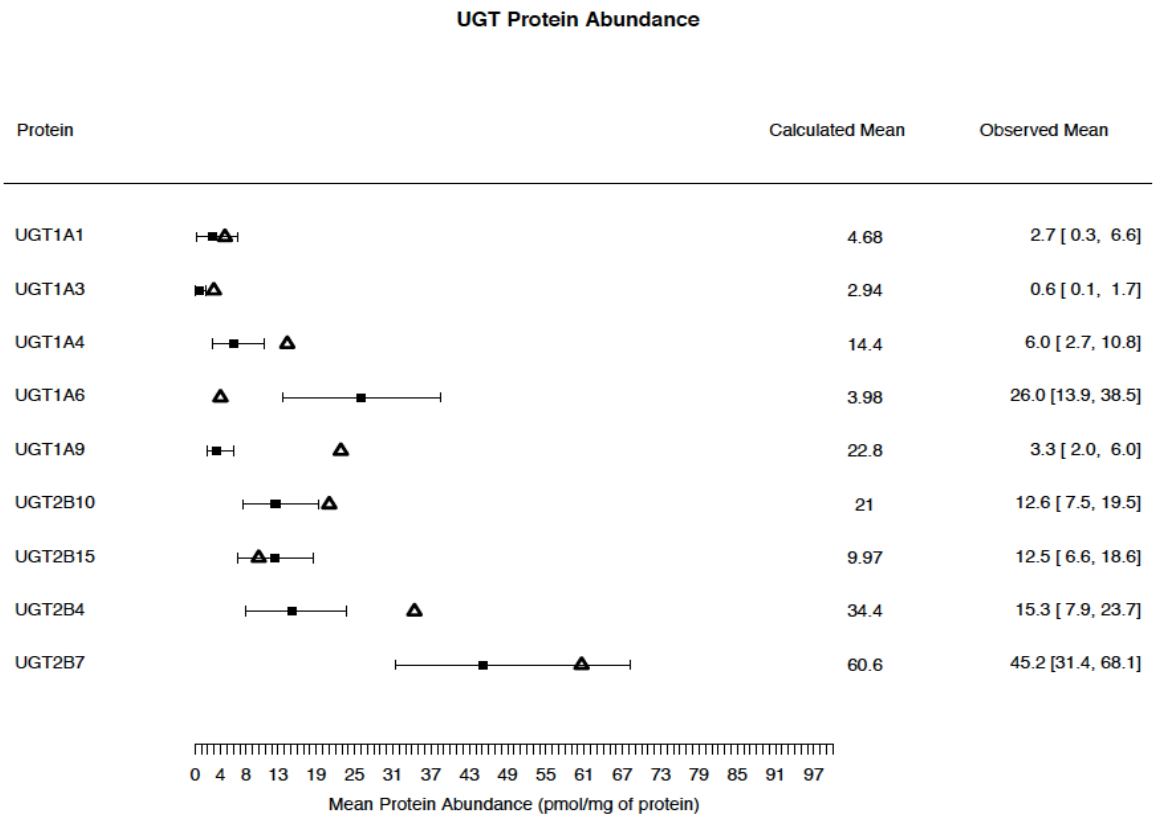

Figure 2.12: Comparison of UGT proteins using SWATH-MS DIA-TPA data with published data (Weisnewski et. al, 2016) based on DDA-TPA approach.

The square in black represent the average of human hepatocytes (n=6) enzymes with min and max range. The triangle represents the average of human hepatocytes from current study.

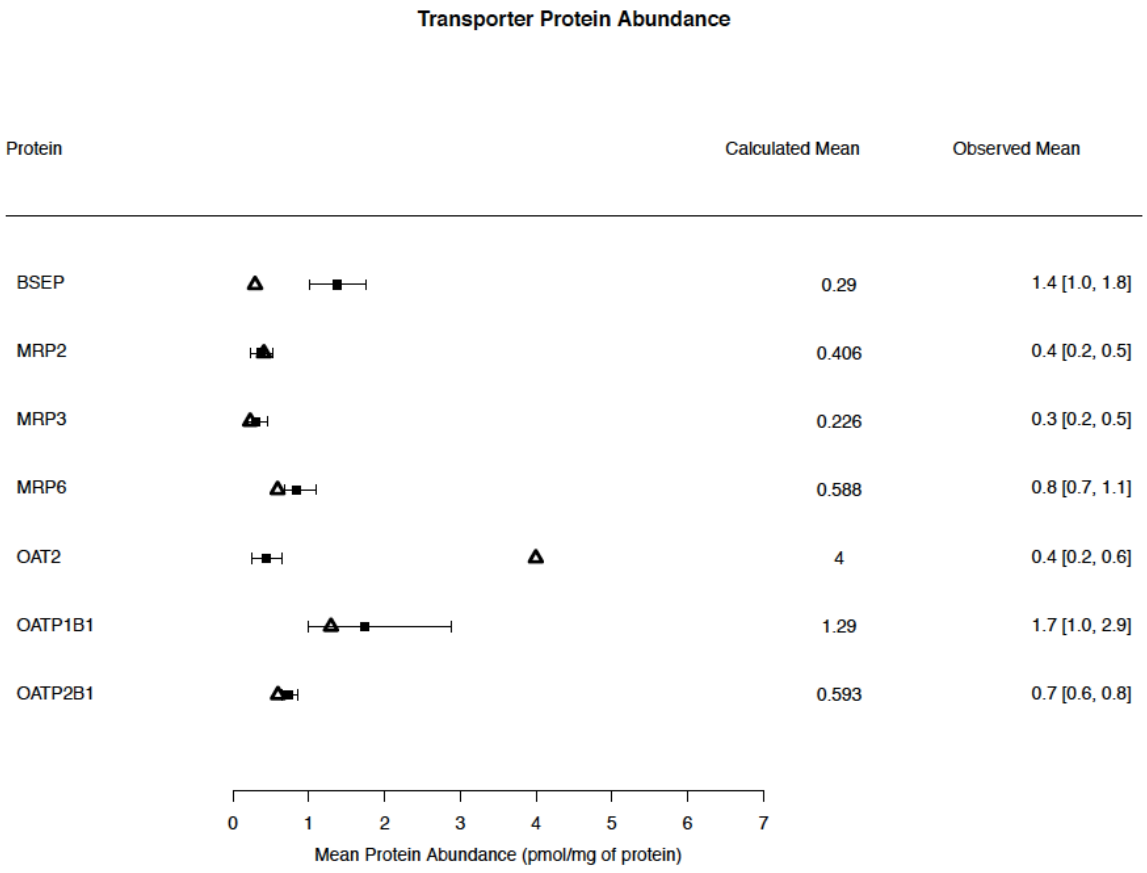

Figure 2.13: Comparison of transport proteins using SWATH-MS DIA-TPA data with published data (Weisnewski et. al, 2016) based on DDA-TPA approach.

The square in black represent the average of human hepatocytes (n=6) enzymes with min and max range. The triangle represents the average of human hepatocytes from current study.

Table 2.1: Demographic information of the cryopreserved human hepatocytes.

Product #	Lot #	Viability	Yield (million cells)	Age (years)	Race	Sex	BMI
M00995-P	**LQG**	81%	8.98	32	WH	M	23.5
M00995-P	**AQL**	91%	8.34	63	WH	M	26.1
F00995-P	**VHB**	86%	8.30	57	WH	F	26.4

Table 2.2: Global hepatic protein in HepaRG, HepG2, Huh7 cell lines and human hepatocytes identified using Spectronaut

	Human Hepatocytes			***Human Liver Tissue***			***HepaRG***			***HepG2***			***Huh7***		
	1	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
Precursor	34760	36127	36148	43914	32120	46156	29656	33577	32400	24995	27818	28952	33831	32760	30635
Peptides	24791	25499	25630	30040	31460	31412	22489	25041	24283	19486	21423	21978	25474	24726	23277
Proteins	2966	3028	3043	3335	3382	3,371	2904	3132	3057	2808	2967	2990	3249	3203	3088
Proteins represented in all replicates	2717			3083			2715			2578			2874		

Table 2.3: The proteins expressed in gluconeogenesis/glycolysis detected using SWATH-MS.

The average of the triplicates is tabulated and expressed in pmol/mg of protein. NA denote the protein expression below detectable limits.

Gene Name	Location	Function	cHH (pmol/ mg of protein)	HLT (pmol/ mg of protein)	HepaRG (pmol/ mg of protein)	HepG2 (pmol/ mg of protein)	Huh7 (pmol/ mg of protein)
ALDOA	Cytoplasm	Gluconeogenesis/ Glycolytic process	1.2	5.2	15.0	40.5	25.4
ENO1	Mitochondrion	Cell membrane, cytoplasm and nucleus	13.6	60.8	87.1	145.2	55.2
FBP1	Cytoplasm	Gluconeogenesis	5.7	26.3	NA	NA	NA
FBP2	Cytoplasm	Gluconeogenesis	2.4	19.0	NA	NA	NA
G6PC	Cytoplasm	Gluconeogenesis	1.1	0.9	0.3	NA	0.7
GAPDH	Nucleus, cytoskeleton and cytosol	Gluconeogenesis/ Glycolytic process	89.7	224.2	155.2	432.2	71.2
GPI	Cytoplasm	Glycolysis	1.1	6.9	6.8	7.6	2.0
HK1	Mitochondrion membrane protein	Glycolysis	5.2	2.5	2.2	0.8	1.0
LDHA	Cytoplasm	Pyruvate metabolism	38.0	204.8	175.8	351.6	122.3
LDHB	Cytoplasm	Pyruvate metabolism	1.0	2.6	NA	0.6	NA
PC	Mitochondria	Gluconeogenesis	20.7	14.6	3.2	2.7	3.0
PCK1	Cytoplasm	Gluconeogenesis	1.8	2.8	16.9	NA	0.3
PCK2	Mitochondria	Gluconeogenesis	84.5	21.5	13.8	3.3	17.7
PFKL	Cytoplasm	Glycolysis	0.7	1.5	0.4	1.5	0.5
PGAM1	Nucleus	Glycolysis	4.3	20.8	13.4	43.3	5.7
PGK1	Cytosol	Glycolysis	6.9	28.0	36.7	66.6	14.5

PKLR	Cytosol	Glycolysis	4.2	6.9	32.0	80.9	12.0
PKM	Nucleus	Glycolysis	1.6	2.3	72.9	173.7	22.8
TPI1	Cytoplasm	Gluconeogenesis	14.6	64.0	25.1	126.2	22.5

Table 2.4: The proteins expressed in lipid metabolism detected using SWATH-MS.

The average of the triplicates is tabulated and expressed in pmol/mg of protein. NA - Protein expression below detectable limits.

Gene Name	Location	Function	cHH (pmol/ mg of protein)	HLT (pmol/ mg of protein)	HepaRG (pmol/ mg of protein)	HepG2 (pmol/ mg of protein)	Huh7 (pmol/ mg of protein)
ACSL1	Peroxisome, ER and mitochondria	Acyl-CoA Ligase	38.2	35.3	27.6	0.4	1.0
ACSL3	Peroxisome, ER and mitochondria	Acyl-CoA Ligase	0.8	1.8	3.4	6.1	11.1
ACSL5	Peroxisome, ER and mitochondria	Acyl-CoA Ligase	7.7	6.5	6.3	3.9	15.3
ACAA1	Peroxisome	Peroxisomal ***β***-oxidation	47.2	24.5	15.3	9.6	4.4
ACOX1	Peroxisome	Peroxisomal ***β***-oxidation	3.2	1.9	1.6	1.2	1.5
ACOX2	Peroxisome	Peroxisomal ***β***-oxidation	7.4	3.5	0.8	0.9	0.8
ACOX3	Peroxisome	Peroxisomal ***β***-oxidation	0.7	0.6	0.7	0.4	0.3
EHHADH	Peroxisome	Peroxisomal ***β***-oxidation	17.1	10.8	8.5	0.6	0.8
HSD17B4	Peroxisome	Peroxisomal ***β***-oxidation	57.7	42.6	34.8	10.6	12.9
CPT1A	Mitochondria	Carnitine Shuttle	1.6	1.4	1.6	0.5	2.3
CPT2	Mitochondria	Carnitine Shuttle	6.6	6.6	4.5	1.5	1.5
SLC25A20	Mitochondria	Carnitine Shuttle	9.8	4.9	4.4	2.0	1.6
ACAA2	Mitochondria	Mitochondrial ***β***-oxidation	125.6	79.3	30.7	7.2	7.8
ACADL	Mitochondria	Mitochondrial ***β***-oxidation	1.0	0.4	1.5	NA	NA
ACADM	Mitochondria	Mitochondrial ***β***-oxidation	28.3	20.4	13.5	2.8	5.3
ACADS	Mitochondria	Mitochondrial ***β***-oxidation	16.1	13.4	0.9	1.0	2.2
ACADSB	Mitochondria	Mitochondrial ***β***-oxidation	20.6	12.4	2.8	2.1	9.9
ACADVL	Mitochondria	Mitochondrial ***β***-oxidation	9.8	10.9	10.7	3.2	2.8
ECHDC1	Mitochondria	Mitochondrial ***β***-oxidation	1.4	2.6	1.9	0.9	1.4
ECHDC2	Mitochondria	Mitochondrial ***β***-oxidation	27.0	14.2	1.4	NA	NA

ECHDC3	Mitochondria	Mitochondrial β-oxidation	8.5	3.8	9.3	0.8	NA
ECHS1	Mitochondria	Mitochondrial β-oxidation	181.5	110.4	35.2	17.9	13.5
HADH	Mitochondria	Mitochondrial β-oxidation	48.6	40.4	23.1	5.2	9.5
HADHA	Mitochondria	Mitochondrial β-oxidation	28.4	23.4	20.4	6.9	17.6
HADHB	Mitochondria	Mitochondrial β-oxidation	15.4	12.6	12.6	3.4	10.4

Table 2.5: Top 5 differentially regulated pathway in HepaRG, Huh7, HepG2 when compared to human hepatocytes.

Rank	HepaRG vs. Human Hepatocytes				HuH7 vs. Human Hepatocytes				HepG2 vs. Human Hepatocytes			
	Pathway name	**#Entities found**	**#Entities total**	**pValue**	**Pathway name**	**#Entities found**	**#Entities total**	**pValue**	**Pathway name**	**#Entities found**	**#Entities total**	**pValue**
1	EIF2 Signaling	99	224	3.94E-55	EIF2 Signaling	125	224	1.14E-78	EIF2 Signaling	106	224	7.19E-66
2	Regulation of eIF4 and p70S6K Signaling	59	157	1.21E-28	Mitochondrial Dysfunction	87	171	1.83E-50	Regulation of eIF4 and p70S6K Signaling	57	157	2.16E-28
3	Remodeling of Epithelial Adherens Junctions	34	69	1.19E-21	Oxidative Phosphorylation	66	109	4.30E-45	Sirtuin Signaling Pathway	75	292	7.26E-26
4	mTOR Signaling	58	210	1.79E-20	Regulation of eIF4 and p70S6K Signaling	78	157	2.35E-44	Mitochondrial Dysfunction	54	171	1.36E-23
5	Sirtuin Signaling Pathway	69	292	5.79E-20	Sirtuin Signaling Pathway	95	292	8.47E-35	mTOR Signaling	59	210	9.59E-23

6. References

1 Liver, E. A. F. T. S. O. T. & Cancer, E. O. F. R. A. T. O. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol* **56**, 908-943, doi:10.1016/j.jhep.2011.12.001 (2012).

2 Wiśniewski, J. R., Vildhede, A., Norén, A. & Artursson, P. In-depth quantitative analysis and comparison of the human hepatocyte and hepatoma cell line HepG2 proteomes. *J Proteomics* **136**, 234-247, doi:10.1016/j.jprot.2016.01.016 (2016).

3 Toyoda, H. *et al.* Changes to hepatocyte ploidy and binuclearity profiles during human chronic viral hepatitis. *Gut* **54**, 297-302, doi:10.1136/gut.2004.043893 (2005).

4 Nagarajan, S. R. *et al.* Lipid and glucose metabolism in hepatocyte cell lines and primary mouse hepatocytes: a comprehensive resource for in vitro studies of hepatic metabolism. *Am J Physiol Endocrinol Metab* **316**, E578-E589, doi:10.1152/ajpendo.00365.2018 (2019).

5 Dave, T., Tilles, A. W. & Vemula, M. A Cell-Based Assay to Investigate Hypolipidemic Effects of Nonalcoholic Fatty Liver Disease Therapeutics. *SLAS Discov* **23**, 274-282, doi:10.1177/2472555217741077 (2018).

6 Chavez-Tapia, N. C., Rosso, N. & Tiribelli, C. Effect of intracellular lipid accumulation in a new model of non-alcoholic fatty liver disease. *BMC Gastroenterol* **12**, 20, doi:10.1186/1471-230X-12-20 (2012).

7 Brown, M. V., Compton, S. A., Milburn, M. V., Lawton, K. A. & Cheatham, B. Metabolomic signatures in lipid-loaded HepaRGs reveal pathways involved in steatotic progression. *Obesity (Silver Spring)* **21**, E561-570, doi:10.1002/oby.20440 (2013).

8 Ma, R. *et al.* Switch of glycolysis to gluconeogenesis by dexamethasone for treatment of hepatocarcinoma. *Nat Commun* **4**, 2508, doi:10.1038/ncomms3508 (2013).

9 Gerets, H. H. *et al.* Characterization of primary human hepatocytes, HepG2 cells, and HepaRG cells at the mRNA level and CYP activity in response to inducers and their predictivity for the detection of human hepatotoxins. *Cell Biol Toxicol* **28**, 69-87, doi:10.1007/s10565-011-9208-4 (2012).

10 Guillouzo, A. *et al.* The human hepatoma HepaRG cells: a highly differentiated model for studies of liver metabolism and toxicity of xenobiotics. *Chem Biol Interact* **168**, 66-73, doi:10.1016/j.cbi.2006.12.003 (2007).

11 Sivertsson, L. *et al.* CYP3A4 catalytic activity is induced in confluent Huh7 hepatoma cells. *Drug Metab Dispos* **38**, 995-1002, doi:10.1124/dmd.110.032367 (2010).

12 Vildhede, A., Wiśniewski, J. R., Norén, A., Karlgren, M. & Artursson, P. Comparative Proteomic Analysis of Human Liver Tissue and Isolated Hepatocytes with a Focus on Proteins Determining Drug Exposure. *J Proteome Res* **14**, 3305-3314, doi:10.1021/acs.jproteome.5b00334 (2015).

13 Gu, X. *et al.* Comparative untargeted proteomic analysis of ADME proteins and tumor antigens for tumor cell lines. *Acta Pharm Sin B* **8**, 252-260, doi:10.1016/j.apsb.2017.10.002 (2018).

14 Shi, J., Wang, X., Lyu, L., Jiang, H. & Zhu, H. J. Comparison of protein expression between human livers and the hepatic cell lines HepG2, Hep3B, and Huh7 using SWATH and MRM-HR proteomics: Focusing on drug-metabolizing enzymes. *Drug Metab Pharmacokinet* **33**, 133-140, doi:10.1016/j.dmpk.2018.03.003 (2018).

15 Huang, Q. *et al.* SWATH enables precise label-free quantification on proteome scale. *Proteomics* **15**, 1215-1223, doi:10.1002/pmic.201400270 (2015).

16 Gillet, L. C. *et al.* Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics* **11**, O111.016717, doi:10.1074/mcp.O111.016717 (2012).

17 Gao, Y. *et al.* Quantitative proteomics by SWATH-MS reveals sophisticated metabolic reprogramming in hepatocellular carcinoma tissues. *Sci Rep* **7**, 45913, doi:10.1038/srep45913 (2017).

18 Xu, M. *et al.* Targeted LC-MS/MS Proteomics-Based Strategy To Characterize in Vitro Models Used in Drug Metabolism and Transport Studies. *Anal Chem*, doi:10.1021/acs.analchem.8b01913 (2018).

19 Achour, B., Al Feteisi, H., Lanucara, F., Rostami-Hodjegan, A. & Barber, J. Global Proteomic Analysis of Human Liver Microsomes: Rapid Characterization and Quantification of Hepatic Drug-Metabolizing Enzymes. *Drug Metab Dispos* **45**, 666-675, doi:10.1124/dmd.116.074732 (2017).

20 Jamwal, R. *et al.* Nonalcoholic Fatty Liver Disease and Diabetes Are Associated with Decreased CYP3A4 Protein Expression and Activity in Human Liver. *Mol Pharm* **15**, 2621-2632, doi:10.1021/acs.molpharmaceut.8b00159 (2018).

21 Jamwal, R. *et al.* Multiplex and Label-Free Relative Quantification Approach for Studying Protein Abundance of Drug Metabolizing Enzymes in Human Liver Microsomes Using SWATH-MS. *J Proteome Res* **16**, 4134-4143, doi:10.1021/acs.jproteome.7b00505 (2017).

22 Prasad, B. *et al.* Interindividual variability in hepatic organic anion-transporting polypeptides and P-glycoprotein (ABCB1) protein expression: quantification by liquid chromatography tandem mass spectroscopy and influence of genotype, age, and sex. *Drug Metab Dispos* **42**, 78-88, doi:10.1124/dmd.113.053819 (2014).

23 Li, C., Zhang, G., Zhao, L., Ma, Z. & Chen, H. Metabolic reprogramming in cancer cells: glycolysis, glutaminolysis, and Bcl-2 proteins as novel therapeutic targets for cancer. *World J Surg Oncol* **14**, 15, doi:10.1186/s12957-016-0769-9 (2016).

24 Fraczek, J., Bolleyn, J., Vanhaecke, T., Rogiers, V. & Vinken, M. Primary hepatocyte cultures for pharmaco-toxicological studies: at the busy crossroad of various anti-dedifferentiation strategies. *Arch Toxicol* **87**, 577-610, doi:10.1007/s00204-012-0983-3 (2013).

25 Younossi, Z. M. *et al.* Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology* **64**, 73-84, doi:10.1002/hep.28431 (2016).

26 Gripon, P. *et al.* Infection of a human hepatoma cell line by hepatitis B virus. *Proc Natl Acad Sci U S A* **99**, 15655-15660, doi:10.1073/pnas.232137699 (2002).

27 Parent, R., Marion, M. J., Furio, L., Trépo, C. & Petit, M. A. Origin and characterization of a human bipotent liver progenitor cell line. *Gastroenterology* **126**, 1147-1156 (2004).

28 Hart, S. N. *et al.* A comparison of whole genome gene expression profiles of HepaRG cells and HepG2 cells to primary human hepatocytes and human liver tissues. *Drug Metab Dispos* **38**, 988-994, doi:10.1124/dmd.109.031831 (2010).

29 Jennen, D. G. *et al.* Comparison of HepG2 and HepaRG by whole-genome gene expression analysis for the purpose of chemical hazard identification. *Toxicol Sci* **115**, 66-79, doi:10.1093/toxsci/kfq026 (2010).

30 Schwanhäusser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473**, 337-342, doi:10.1038/nature10098 (2011).

31 Slany, A. *et al.* Cell characterization by proteome profiling applied to primary hepatocytes and hepatocyte cell lines Hep-G2 and Hep-3B. *J Proteome Res* **9**, 6-21, doi:10.1021/pr900057t (2010).

32 Sison-Young, R. L. *et al.* Comparative Proteomic Characterization of 4 Human Liver-Derived Single Cell Culture Models Reveals Significant Variation in the Capacity for Drug Disposition, Bioactivation, and Detoxication. *Toxicol Sci* **147**, 412-424, doi:10.1093/toxsci/kfv136 (2015).

33 Ni, R., Leo, M. A., Zhao, J. & Lieber, C. S. Toxicity of beta-carotene and its exacerbation by acetaldehyde in HepG2 cells. *Alcohol Alcohol* **36**, 281-285, doi:10.1093/alcalc/36.4.281 (2001).

34 Lu, D., Wang, S., Xie, Q., Guo, L. & Wu, B. Transcriptional Regulation of Human UDP-Glucuronosyltransferase 2B10 by Farnesoid X Receptor in Human Hepatoma HepG2 Cells. *Mol Pharm* **14**, 2899-2907, doi:10.1021/acs.molpharmaceut.6b01103 (2017).

In Preparation for submission to Liver International

MANUSCRIPT 3

Transcriptomic and Proteomics Analysis of a Lipid-Loaded HepaRG Model for Steatosis Reveals Altered Regulation in Lipid and Xenobiotic Metabolism

Anitha Saravanakumar [1], Cassandra A. Tierney [2], Wen He [2], Rohitash Jamwal [1], Benjamin Barlock [1], Xin Bush [1], Jillian G. Johnson [2], David Rodrigues [3], Fatemeh Akhlaghi [1*]

[1]*Clinical Pharmacokinetics Research Laboratory, Department of Biomedical and Pharmaceutical Sciences, University of Rhode Island, Kingston, RI, USA*

[2] *Clinical Pharmacogenomics Laboratory, Pfizer Worldwide Research, Development and Medical, Groton, CT, USA*

[3]*ADME Sciences, Medicine Design, Pfizer Worldwide Research, Development and Medical, Groton, CT, USA*

Corresponding author at:

Fatemeh Akhlaghi, Ph.D.

Clinical Pharmacokinetics Research Laboratory

Department of Biomedical and Pharmaceutical Sciences

The University of Rhode Island, 495A College of Pharmacy, 7 Greenhouse Road,

Kingston, RI 02881, United States.

Email address: fatemeh@uri.edu

Abbreviations

4HNE - 4-Hydroxynonenal

ACSL1 - Acyl- CoA Synthetase Long Chain Family Member 1

ACSL4 - Acyl- CoA Synthetase Long Chain Family Member 4

ADME - Absorption, Distribution, Metabolism, and Excretion

AKR1B10 - Aldo–Keto Reductase 1B10

APOE - Apolipoprotein E

BSA – Bovine Serum Albumin

CAR - Constitutive Androstane Receptor

CCL2 - C-C Motif Chemokine Ligand 2

CYP – Cytochrome P450

DE - Differentially Expressed

DGAT1 - Diacylglycerol O-Acyltransferase 1

DME - Drug Metabolizing Enzyme

DMET - Drug Metabolizing Enzyme and Transporters

DMSO - Dimethyl Sulfoxide

EDTA - Ethylene- Diamine-Tetracetic Acid

EMEM - Eagle Minimal Essential Medium

ENHO - Energy Homeostasis Associated

FABP4 - Fatty Acid Binding Protein 4

FA – Fatty Acid

FAF - Fatty Acid Free

G6PC - Glucose-6-Phosphatase Catalytic Subunit

GEO - Gene Expression Omnibus

GNMT - Glycine N-Methyltransferase

GPS2 - G Protein Pathway Suppressor 2

GPT - Glutamic-Pyruvic Transaminase

HNF4α - Hepatocyte Nuclear Factor 4α

HSD17B13 - Hydroxysteroid 17-Beta Dehydrogenase 13

IGFBP1 - Insulin Like Growth Factor Binding Protein 1

IL6 - Interleukin 6

KLF2 - Kruppel Like Factor 2

LBP - Lipopolysaccharide Binding Protein

Lds - Lipid Droplets

LIPC - Lipase C, Hepatic Type

LPL - Lipoprotein Lipase

MCD - Methionine Choline Deficient

MDA - Malondialdehyde

MIGA2 - Mitoguardin 2

MOGAT3 - Monoacylglycerol O-Acyltransferase 3

NAFLD - Non- Alcoholic Fatty Liver Disease

NASH - Non-Alcoholic Steatohepatitis

NCBI - National Center Biotechnology Information

NCE - New Chemical Entities

NEIL1 - Nei Like DNA Glycosylase 1

ORO - Oil Red O Staining

P450 – Cytochrome P450

PBS - Phosphate- Buffered Saline

PCR - Polymerase Chain Reaction

PHH - Primary Human Hepatocytes

PLIN2 - Perilipin 2

PPARα - Peroxisome proliferator-activated receptor

PXR - Pregnane X Receptor

RGN - Regucalcin

RIPA - Radio Immunoprecipitation Assay Buffer

RNA - Ribonucleic Acid

ROS - Reactive Oxygen Species

RT - Room Temperature

SCG5 - Secretogranin V

SIRT4 - Sirtuin 4

SLC2A4 - Solute Carrier Family 2 Member 4

SOD2 - Superoxide Dismutase 2

TG – Triglyceride

TCA - Tricarboxylic Acid Cycle

TLR2 - Toll Like Receptor 2

TM6SF2 - Transmembrane 6 Superfamily Members 2

TXNIP - Thioredoxin Interacting Protein

log2FC – Log fold change

Abstract

Background & Aims. Hepatic lipid accumulation (steatosis) is an early sign of a spectrum of Non-alcoholic Fatty Liver Disease (NAFLD) that precedes fibrosis and cirrhosis. To manage NAFLD and related co-morbidities, patients are administered an array of pharmacological agents. Therefore, understanding of the effect of NAFLD on drug disposition is warranted. Using a HepaRG model, we aimed to mimic steatosis *in vitro*, and to examine its effects on drug-metabolizing enzymes (DME) and transporters.

Methods. HepaRG cells, differentiated in-house, were exposed to a mixture of saturated and unsaturated fatty acids (1:2 ratio of 0.5 mM palmitate and oleate complex conjugated to BSA for 72h) and were subjected to RNA sequencing and proteomic analyses. Lipid accumulation was ascertained by Oil Red O (ORO) staining and triglyceride (TG) quantitation and cell viability by WST-1 determination.

Results. The treatment condition resulted in ~6-fold increase in TG concentration without reducing cell viability. RNA sequencing and Proteomics (SWATH-MS) of lipid-loaded and control cells identified a total of 393 differentially expressed transcripts (89 up- and 304 down-regulated) and 127 differentially expressed proteins (127 up- and 38 downregulated), respectively. Moreover, lipid loading resulted in significant down-regulation of mRNA transcripts of transcription factors, NR1I2 (-1.18) and HNF4α (-0.55), phase 1 DMEs including CYP1A2 (-3.25), 2B6 (-2.02), 2C8 (-1.48), 2C9 (-2.00),2C19 (-1.32) and 3A4 (-1.87), phase 2 DMEs including UGT1A6 (-0.36) and 2B7 (-1.09), SULT2A1 (-0.75) and 1E1 (-1.41) as well as clinically relevant transporters such as ABCC11 (-1.36), ABCG5 (-1.66), SLC10A1(-1.63), SLCO2B1(-1.49) and SLC10A1 (-1.63). Furthermore, lipid loaded cells significantly upregulated AKR1B10 mRNA (2.17) and protein (0.99) that may regulate lipid as well as xenobiotic metabolism.

Conclusion. Expression of several DME and transporter mRNA was altered in an *in vitro* model of hepatic steatosis which may translate to altered drug disposition in NAFLD patients.

Keywords

Hepatic steatosis, Non-alcoholic Fatty Liver Disease, Cytochrome P450, Drug-Metabolizing Enzymes, Transporters**,** HepaRG.

1. Introduction

Nonalcoholic fatty liver disease (NAFLD) is highly prevalent in the United States as well as other developed counties (Bellentani, 2017). Excess intake of calories and lack of physical activity can cause accumulation of free fatty acids (FAF) and triglyceride (TG) in hepatocytes leading to the development of steatosis (Veteläinen *et al.*, 2007). NAFLD which is associated with metabolic syndrome, includes a spectrum of liver disorders, initiating with simple steatosis, progressing to nonalcoholic steatohepatitis (NASH) and possibly lead to end-stage cirrhosis (Koo, 2013; Cobbina and Akhlaghi, 2017). Furthermore, patients with steatosis are at a greater risk of developing hepatocellular carcinoma (HCC) (Ohata *et al.*, 2003). The benign nature of steatosis is considered as the "first hit" in the progression of NAFLD. The "second-hit" is manifested via reactive oxygen species (ROS) induced stress, which leads to the infiltration of inflammatory factors, hepatocellular ballooning and deposition of collagen fibers in hepatocytes as observed in NASH and fibrosis (Caldwell and Argo, 2010).

Approximately 29% of the patients diagnosed with NASH will develop cirrhosis later in life (Cohen, Horton and Hobbs, 2011). Currently, many pharmaceutical companies are aspiring to develop medications for the treatment of NAFLD (Noureddin, Zhang and Loomba, 2016; Brodosi *et al.*, 2016). Moreover, considering that NAFLD is associated with other co-morbidities including diabetes (Brunt and Tiniakos, 2010), patients are administered several other medications and are at a greater risk of adverse drug reactions. Recently, it was shown that the ADME (Absorption, Distribution, Metabolism and Excretion) properties of midazolam and morphine is altered in NAFLD or NASH (Jamwal *et al.*, 2018; Farrow and Facchini, 2013). Hence, a comprehensive understanding of the differential expression of Drug-Metabolizing Enzyme and Transporters (DMET) in NAFLD is warranted to predict the PK properties of pharmaceutical agents.

To study steatosis, it is crucial to identify an appropriate disease model. Significant differences exist in the manifestation of this disease between humans and rodents. Diet-induced mouse models fail to recapitulate the full spectrum of NAFLD manifestation in human (Gómez-Lechón, Jover and Donato, 2009; Merrell and Cherrington, 2011). Feeding high-fat diet (HFD) to mice generate simple steatosis; however, male C57BL/6 mice fail to progress to NASH even in the presence of a high fructose-rich diet (Takahashi, Soejima and Fukusato, 2012). Whereas, a methionine-choline-deficient (MCD) diet induces a phenotype of severe NASH without metabolic syndrome and obesity features [16]. In addition, genetically modified mouse models including *ob/ob* or *db/db* often fail to recapitulate the full spectrum of the disease (Larter and Yeh, 2008). Collectively, the disparity between rodent models and human with respect to the regulation and tissue distribution of DMETs (Prueksaritanont, 2010) (Martignoni, Groothuis and de Kanter, 2006) limit the usefulness of data obtained from animal models of NAFLD/NASH.

Use of primary human hepatocytes (PHH) or cryopreserved hepatocytes is considered the gold standard for studying DMET *in vitro* (Lecluyse and Alexandre, 2010). In addition, FA-loading has been previously attempted with primary human and rat hepatocytes using static and dynamic chip-based systems (Donato *et al.*, 2006; Kostrzewski *et al.*, 2017; Bulutoglu *et al.*, 2019). However, the availability of early-stage biopsy samples, donor-to-donor variability, cost involved in model development and limited time in culture before the onset of dedifferentiation diminishes the utility of PHH as an effective tool for disease model development. In this regard, the HepaRG cell line is a promising *in vitro* model presenting a more robust DMET gene expression pattern than HepG2 and Huh7 cell lines (Gerets *et al.*, 2012). HepaRG cells are an immortalized human hepatocellular carcinoma cell line comprised of a mixture of both hepatocyte-like and biliary-like cells (Guillouzo *et al.*, 2007). Furthermore, they preserve their phenotype for over 2 weeks upon

differentiation (Aninat *et al.*, 2006). These attributes make HepaRG cells a suitable cell line for the development of a disease model and its utilization for studying therapeutic modalities.

We aim to characterize the molecular basis of NAFLD using a HepaRG model aiming to understand the regulation of DMET in the disease state. Steatosis was induced by treating HepaRG cells with a mixture of saturated and unsaturated fatty acids (FA). To obtain steatotic global gene expression profiles, we carried out RNA sequencing (RNA-Seq) on HepaRG cells treated with FA or control. Subsequently, the proteomics signature, altered activity of cytochrome P450 (CYP) 3A4 and CYP2E1 were recorded in this model of steatosis.

2. Material and Methods

2.1 Cell Culture Model

HepaRG cells were purchased from Biopredic International (through licensing agreement from Inserm, France) and were grown and differentiated in-house according to manufacturer's instructions. The incubation was carried out in a at 37 °C with 5% CO2 (New Brunswick Galaxy® incubator CO2, Model# 170R, New Brunswick, NJ). Following 14 days growth and 14 days differentiation phases, the cells were seeded in a 24 well plate, and were treated with basal medium containing Williams E with Glutamax-I, 100 IU/ml of penicillin, 100 µg/ml of streptomycin, 4 µg/ml of bovine insulin and 5×10^{-5} M of hydrocortisone hemisuccinate for 5 days.

To induce steatosis, HepaRG cells were treated with FA (1:2 palmitate: oleate) conjugated to 30% essentially fatty acid-free (FAF) BSA solution (Sigma-Aldrich, Catalog# A9576, Saint Louis, MO) made up in Williams E media (Thermo-Fisher Scientific, Catalog# 12551-032, Waltham, MA). The control cells were treated with 30% essentially FAF BSA solution alone. The cells were treated for 24h, 48h and 72h with media replenished after 48h (Brown *et al.*, 2013).

2.2 Oil Red O Staining

HepaRG cells treated with 0.5 mM FA or control for 24h, 48h and 72h. Cells were washed twice with 1X PBS and fixed with 4% paraformaldehyde (Sigma-Aldrich, Catalog# 100496, Saint Louis, MO) for 30 min in room temperature (RT). Following this, 500 µl of working Oil Red O (ORO) solution (3 parts of stock ORO and 2 parts of dH_2O) (Sigma-Aldrich, Catalog# O0625, Saint Louis, MO) was added and incubated at RT for 20 min. The cells were washed 3X with distilled water (dH_2O), and 500 µL of hematoxylin (Sigma-Aldrich, Catalog# H3136, Saint Louis, MO) was added for 1 min to stain the nuclei and their image was obtained using EVOS cells imaging systems (Life Technologies, Model#

M7000, Foster City, CA). The extent of neutral lipid accumulation was quantified after extraction in 100% isopropanol, and the absorbance was measured at 492 nm using a SpectraMax plate reader (Molecular Devices, Model# M2, Sunnyvale, CA).

2.3 WST-1 cell viability

Cells were treated with 0.1-2 mM BSA conjugated with FA (1:2 palmitate: oleate) for 72h in a 96 well format. The positive control used for this experiment was clotrimazole (100 µM). Following treatment, WST-1 reagent (Roche Diagnostics, Catalog# 11644807001, Mannheim, Germany) was diluted 100X times in HepaRG basal media. The media was added to cells and incubated in 37 °C for 4h. The supernatant was collected and analyzed using the SpectraMax M2 plate reader at 440 and 600 nm wavelengths.

2.4 TG Accumulation

Approximately 10 million cells (HepaRG control vs. FA) were rinsed with PBS and scraped off the flask using a rubber policeman cell scraper. Cells were resuspended in PBS and lysed by sonication twenty times with one-second bursts. The samples were centrifuged at 10,000 xg for 10 minutes at 4 °C, and the supernatant was collected and assayed for TG content using a TG colorimetric assay kit (Cayman Chemicals, Catalog# 10010303, Ann Arbor, MI) according to manufacturer's protocol. The optical density was read at 540nm on SpectraMax M2 plate reader.

2.5 Cholesterol Assay

One million HepaRG cells (0.5 mM FA treatment vs. control) were trypsinized, and cell pellets were lysed with 200 µL of chloroform: isopropanol: NP-40 (7:11:0.1) in a micro-homogenizer. The cell lysate was centrifuged at 15,000 x g for 10 min. The supernatant was isolated and air dried at 50 °C to remove the chloroform. Samples were placed in a vacuum for 30 minutes to remove the trace amounts of organic solvent. Dried lipids were

mixed in 200 µL of 1X assay diluent with sonication and vortex mixed until dissolved. The extracted samples were assayed using a total cholesterol calorimetric assay kit (Cell Biolabs Inc., Catalog# STA-384, San Diego, CA).

2.6 Assessment of Oxidative Stress

Differentiated HepaRG cells (10 million cells) were treated with BSA or FA treatment (0.5 mM) for 72 hours in triplicates. Oxidative stress was analyzed as a measure of lipid peroxidation to Malondialdehyde (MDA). The samples were quantified using OxiSelect™ TBARS Assay Kit (Cell Biolabs Inc., Catalog# STA-330, San Diego, CA). Following this, samples were analyzed at 542 nm using SpectraMax M2 plate reader.

2.7 Insulin-stimulated phosphorylation of Protein Kinase B (Akt)

Following 72h of FA treatment, HepaRG cells were placed in serum-free EMEM media for 12 h followed by stimulation with recombinant human insulin (Sigma-Aldrich, Catalog# 1376497001, Saint Louis, MO) at 1, 10 or 100 nM for 10 min at 37 °C. Samples were rinsed in PBS and lysed with RIPA buffer (150mM NaCl, 50mM Tris, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 1mM EDTA) supplemented with a complete protease inhibitor cocktail (Sigma-Aldrich, Catalog# CO-RO, Saint Louis, MO) and Halt Protease Inhibitor Cocktail (Thermo Fisher Scientific, Catalog# 78430, Waltham, MA) for 30 min at 4 °C. The lysate was run on a sandwich ELISA kit that detects both pAkt (Ser473) and total Akt (Abcam, Catalog#126433, Cambridge, MA). The assay was performed according to manufacturer's protocol and the results were expressed as a ratio of phospho-Akt to total Akt.

2.8 RNA extraction and sequencing

Total RNA was extracted using RNeasy Mini Kit (Qiagen, Valencia, CA) from differentiated HepaRG cells (1 million cells per well) plated in a 6-well plate (0.5 mM FA treatment vs. control for 72h). The contents of each well were considered as a replicate

for RNA extraction and subsequent sequencing. The ratio of absorbance at 260 nm and 280 nm was used to assess the purity of RNA using Nanodrop (Thermo Fisher Scientific, Catalog #ND-ONE-W, Waltham, MA). Samples with a 260/280 ratio of ~1.8 and RNA Integrity Number (RIN) > 8 were selected for further analysis. RNA integrity and quantity were measured using Agilent 2100 Bioanalyzer System (Agilent Technologies, Santa Clara, CA). Total RNA input for library preparation was assessed using the Qubit RNA BR assay (ThermoFisher Scientific catalog# Q10210). RNA sequencing was performed on 100ng of oligo(dT) purified poly(A)+ mRNA. A standard TruSeq mRNA library was constructed using TruSeq® Stranded mRNA NeoPrep™ Kit v2 (Illumina, Catalog # NP-202-1001, San Diego, CA). The library was sequenced by Illumina NextSeq 500 using a paired-end run (2×75 bases). The fastq files were generated using BaseSpace via bcl2fastq.

2.9 RNA-seq data analysis

Sequencing data have been uploaded in the Gene Expression Omnibus (GEO) database that can be viewed with the accession number of GSE122151. The data analysis and visualization were carried out using QuickRNASeq, an integrated pipeline used in the analysis for large-scale RNA-Seq data (Zhao *et al.*, 2016). Read mapping and counting was performed using STAR and featureCounts. Following this, RSeQC package was included to employ all the QC metrics and to remove outliers; whereas VarScan, platform-independent software tool was included for variant detection in RNA-Seq data. A generalized linear model was used instead of a pairwise comparison approach to account for the changes between treatment and control according to the R user's guide to determine differentially expressed genes. We used log2 FC > 0.58 as a cut-off combined with exclusion criteria with False Discovery Rate (FDR) < 0.01 and p-value < 0.05, to define the differentially expressed genes.

2.10 SWATH-MS acquisition and data analysis

Cell homogenate samples were analyzed using SWATH-MS based spectra which were acquired for mass range m/z 400−1100 Da within SWATH window width of 10 m/z resulting in 70 overlapping mass windows per cycle, as described in the previous literature (Jamwal *et al.*, 2017). The DIA data was searched against Swiss-Prot identifiers (October 2016) and analyzed using Spectronaut™ Pulsar software (version 11.0, Biognosys AG, Schlieren, Switzerland) was used to obtain MS2 signal intensities of fragment ions from DIA data using default settings with a few modifications that includes specific trypsin/P digestion with a minimum of 7 amino acids, maximum of 52 amino acids and 2 missed cleavages; fixed modifications included carbamidomethyl (C) and variable modifications included acetyl (protein N-terminus) and oxidation (M). Raw data files and search results are available at Japan Proteome Standard Repository (jPOSTrepo JPST000372, ProteomeXchange PXD008593). The SWATH-MS DIA data was matched with the reference spectral library generated from the DDA of pooled human liver samples using Protein Pilot. The DIA data was then normalized and analyzed using "Total protein approach" (TPA), as shown below (1).

$$\text{Protein(i)} = \frac{MS\ signal(i)}{Total\ MS\ signal \times \text{Molecular mass(i)}} \times 10^9 \qquad \ldots(1)$$

The absolute protein abundance of each group was represented as pmol/mg of protein.

2.11 Gene Expression Validation

Real-Time PCR. Real-time PCR for human CYP mRNA levels was performed using a ViiA7 Instrument (Thermo Fisher Scientific, Catalog # 4453534, Waltham, MA) and TaqMan assays used are as follows: CYP3A4 (Hs00604506_m1); CYP2E1 (Hs00559367); NR1I3 (Hs00901571_m1); NR1I2 (Hs01114267); GAPDH

(Hs99999905_m1); 18 S (4332641); and β-actin (Hs99999903) (Thermo Fisher Scientific, Waltham, MA). GAPDH, 18 S rRNA and β-actin were used as housekeeping gene controls to normalize the data. The reaction mixture (20 µl/well) contained 30 ng of cDNA (2 µl), 10× Taqman Universal Master Mix (10 µl), Taqman primers (1 µl) and RNase free water (7 µl). The thermal cycle conditions were identical for all the genes analyzed. The qPCR steps include 50 °C for 2 min followed by denaturation at 95 °C for 10 min, 40 cycles of 95 °C for 15 s and 60 °C for 1 min. Fold change was calculated by the formula: Fold = $2^{-\Delta\Delta Ct}$, where ΔCt represents the difference in cycle threshold number between the target gene and the housekeeping gene, whereas ΔΔCt represents the difference between control and fatty acid treatment groups. mRNA samples from individual wells treated with FA treatment or control conditions were analyzed in triplicate, and data were analyzed using Data Assist software (Thermo Fisher Scientific, v3.01, Waltham, MA).

2.12 Characterization of Enzyme Activity

HepaRG cells grown on twenty-four-well plates of were used for activity measurements. After 72 h in culture, treatment media was removed, and the cells were washed twice with 1xPBS; after that, 230 µL of medium was added to each well. The medium consisted of Williams' medium E containing different concentrations of CYP 3A4 substrate, midazolam (0.1, 0.5, 1, 3, 10, 30 µM) (TRC, Catalog# KIT0040, North York, ON, Canada) or CYP2E1 substrate, chlorzoxazone (500, 750, 1000 µM) (TRC, Catalog# C428700, North York, ON, Canada), and incubated for 1h or 8h respectively. The compounds were dissolved in DMSO and then added to the media. The final concentration of DMSO was maintained at 0.1%, and samples were analyzed in triplicates. The reaction was quenched by equal volume of the internal standard containing ice-cold acetonitrile and 0.1% formic acid. The samples were analyzed for the CYP3A4 metabolite 1-hydroxymidazolam or CYP2E1 metabolite 6-hydroxychlorzoxazone (TRC, Catalog#

H825120 North York, ON, Canada). [2H4] hydroxymidazolam (Catalog# H948423) and [2H2] hydroxychlorzoxazone (TRC, Catalog# H825123 North York, ON, Canada) were used as internal standards respectively.

The mass spectrometry parameters were optimized for each analyte (Dostalek *et al.*, 2010). The multiple reaction monitoring transitions chosen were m/z 342 → 203 for 1-hydroxymidazolam and m/z 346 → 203 for d4-hydroxymidazolam. The assay was linear from 1 to 25 ng/mL ($R^2 > 0.99$). The transition chosen for hydroxychlorzoxazone was m/z 183.8 → 119.8 and for d2 hydroxychlorzoxaxone was m/z 187.6 → 121.9. This assay also showed linearity 9 to 925 ng/mL ($R^2 > 0.95$) and precision.

2.13 Data and statistical analyses

RNA-Seq data were analyzed using IPA (Ingenuity Pathway Analysis) software for studying the disease and other canonical pathways, utilizing the generalized pathways that represent common properties of a signaling module using the KEGG database (Qiagen, Foster City, CA). Differential gene expression of FA treated cells was normalized to the control and reported as log2 of the fold change (log2FC). HemI (Heatmap Illustrator, version 1.0) was used to generate the heatmap of differentially expressed genes. In most cases, data were averaged for each experiment performed as triplicates and data were presented as mean ± SEM relative to BSA for each condition. Unpaired Student's t-test analyzed column statistics between control and FA treated groups. Data were analyzed, and graphs were presented using Prism software (GraphPad).

3 Results

3.1 Development of the steatosis model with HepaRG cell line

Upon treatment, the neutral lipid accumulation was visualized at 24h, 48h and 72h utilizing ORO stain (Fig 3.2-I). There was an increasing trend in the lipid accumulation, with 72 h post-treatment showing (Fig 3.2-II) the highest accumulation of lipids. To further define the treatment concentration, varying range of FA (0.1 - 2.0 mM) at 72 h, were assessed for changes in the cell viability. As shown in Fig 3.2-III, 0.5 mM of FA treatment, did not induce any detectable toxicity to the cells and hence was used for further characterization.

3.2 Validation of FA treated HepaRG cells as a model for steatosis

Increased total TG and cholesterol levels were observed under FA treatment conditions. There was a 6-fold increase in triglyceride levels in FA treatment vs. control (Fig. 3.3A). Cholesterol concentration showed ~1.3-fold increase in the FA treatment group compared to control (Fig. 3.3B). To assess the effect of FA treatment on hepatocellular oxidative stress, the mean level of malondialdehyde (MDA) was measured (Fig. 3.3C). However, there was no significant difference between the FA treatment and control cells. This observation suggests that FA treatment may not induce oxidative stress.

Insulin signaling, which is often hindered in steatosis, was assessed in FA treated HepaRG cells. Total Akt levels remained constant for all treatment conditions. At every concentration increment of insulin, a significant reduction in the levels of phospho-Akt was noted within the FA treatment group as compared with control (Fig. 3.3D). This finding shows reduced insulin signaling in FA treated HepaRG cells.

In addition, we compared the transcriptomics data with a clinical study (Starmann *et al.*, 2012) that reported the hepatic gene expression difference between liver steatosis (n=30) and control (n=18). Of the top 100 differentially expressed genes, 24 genes overlapped with our dataset and are summarized in Table 3.1. The genes ACSL4/5, CAT, PNPLA3, ATF3, GSTM5, and NR0B2, were previously reported (Stepanova *et al.*, 2010; Younossi *et al.*, 2005) to show a positive correlation with NAFLD.

3.3 Pathway analysis of differentially regulated genes upon FA treatment vs. control

IPA was used to identify significant pathways affected by differential regulation of genes from FA treatment of HepaRG cells. To further examine the model, we used the differentially expressed (DE) genes to predict the disease state. Two hundred seventy-nine canonical pathways were significantly altered in the dataset of 393 DE genes (Fig. 3.4A). The top twenty pathways are listed in Table 3.2. It is important to interpret this information with caution as the dysregulated canonical pathways were mostly represented by a small fraction of differentially regulated genes (less than 20), in such cases the pathway was not considered significantly dysregulated. Amongst the enriched pathways, the pathways that were dysregulated upon FA treatment, include FXR/RXR, LXR/RXR activation, LPS/IL-1 mediated inhibition of RXR, xenobiotic metabolism, coagulation system and acute phase response signaling (Fig. 3.4C). Differentially expressed genes were also grouped to predict the disease state with the IPA software (Table 3.3). It is interesting to note the top-scoring and statistically significant ($p<0.05$) predictions in the category of metabolic diseases was that of hepatic steatosis, microvesicular steatosis, non-alcoholic fatty liver disease and steatohepatitis (Fig. 3.4B).

3.4 Protein expression pattern in FA treatment vs. control

Similar to the gene expression pattern that predicted hepatosteatosis (Fig. 3.4C), significant changes in protein expression that was associated with hepatocellular injury was observed with fatty acid accumulation. Addition of fatty acids induced the proteins PGK1, ALDOA, PGAM1 involved in glycolytic pathways as seen in steatosis, that leads to the formation of glycerol which is the key component in TG formation (Fig. 3.5A). In addition, we noticed a decrease in albumin production and increase in ALT production as seen in hepatocellular injury (Fig. 3.5B). Moreover, increased protein expression of collagen (COL18A1) and TNF-α are indicators of early stage progression to NASH.

3.5 Effect of FA treatment on immune response as well as lipid catabolism, lipid synthesis and transport

FA treatment significantly upregulated the transcription of genes involved in inflammation (Moylan *et al.*, 2014), FA metabolism as well as some DMET mRNA which are represented in Fig. 3.5B. Genes are known to be involved in inflammation and immune response, including, PDK4 (1.54), CD36 (1.38), CD74 (1.02), TNFAIP6 (2.04), PTGFR (0.73) and IL6 (1.28) were upregulated. This data was positively correlated with the induction of protein expression PDK4 (2.44), CD36 (2.11) and TNFAIP6 (1.08) (data not shown). Additionally, mRNA and protein associated with FA oxidation such as ACSL1, ACSL4, ACOT, ACOX4 were observed to be upregulated. In addition, we also observed increased expression mRNA and protein associated with denovo-lipogenesis ACACB, PLIN2, PLIN4 and SCD. In addition, we noticed increased protein expression of CD36 and FABP4 contributing to the increased influx of FA in the hepatocyte-like cells. Collectively, these results support the idea that FA accumulation in HepaRG cells exhibits broad regulatory roles on immune response, lipogenesis, lipid catabolism and transport.

3.6 Effect of FA treatment on DMEs and nuclear receptor expression

We compared the transcriptomes between FA treatment and control HepaRG cells and identified significant changes in genes, suggesting the role of FA accumulation as an initial trigger to the onset of hepatosteatosis (Fig. 3.4C). Furthermore, FA treatment resulted in differential mRNA expression of the phase I (17 different CYP isoforms, 4 different ADH isoforms and FMO5), phase II (2 different GST isoforms, 7 different UGT isoforms, 2 isoforms of SULT), transporter (5 different ABC superfamily, 3 of SLC superfamily) (Fig. 3.7A). Treatment with FA for 72h shows a statistically significant ($p<0.05$) downregulation of some CYP isoforms, namely, CYP3A4, 1A2, 2B6, 2C8, 2C9 and 2C19 with log fold change in the order of -1.87, -3.25, -2.02, -1.48, -2.00 and -1.32 respectively (Table 3.2). Along with this UGT1A9 (-1.39), UGT1A6 (-0.36), SLCO2B1 (-1.49). In terms of the observed protein expression (Fig. 3.7) we noticed a lower trend in CYP2A6, CYP2C8, CYP2C9 and CYP3A5 as well as UGT2A3 and UGT2B7, however, the results were not statistically significant.

Moreover, the expression pattern of nuclear receptors ad transcriptional factors was further investigated. mRNA expression of NR1I2 (-1.18), NR1I3 (-1.96), PPARα (-0.24) and HNF4α (-0.55) were significantly downregulated (Figure 3.7B). In agreement with the RNA-Seq data, qRT-PCR results indicated that expression of the nuclear receptors PXR (NR1I2) and CAR (NR1I3) and CYP3A4 was markedly reduced upon FA treatment (Fig. 3.7C). Furthermore, mRNA and protein of transcriptional regulator, AKR1B10 was positively increased. This suggests that FA accumulation in steatosis induce AKR1B10 that may modulate the lipid metabolism as well as basal expression of drug-metabolizing enzymes (DMEs) in hepatocytes.

3.7 Functional activity of CYP3A4 in control and fatty acid treatment

CYP3A4, the most abundant human hepatic CYP (Saravanakumar *et al.*, 2019) , along with PXR and CAR expression was assessed using qRT-PCR. CYP3A4 enzyme

activity in FA-treated HepaRG was assessed using midazolam as a substrate. Upon FA treatment, the CYP3A4 enzyme activity was significantly reduced. FA-treatment resulted in a 10-fold reduction in CYP3A4 (maximal rate of metabolism, Vmax) activity compared to the control at 10 μM (Fig. 3.8B). Both our RNA-seq, PCR data as well as enzyme activity showed marked downregulation in CYPs and its corresponding nuclear receptors, demonstrating the link between lipid metabolism in steatosis cells with that of xenobiotic metabolism in liver cells.

4 Discussion

We have established an *in-vitro* model for hepatic steatosis and validated this model by various approaches. Utilizing transcriptomic analyses, we have examined altered molecular signaling pathways upon FA accumulation in HepaRG cells. In addition, we collaborated this with proteomics analysis. Our data agree with earlier reports describing the response to FA loading on lipogenesis as well as inflammatory modulators (Rubin *et al.*, 2015; Parafati *et al.*, 2018).

Hepatic steatosis is characterized by increased accumulation of neutral lipids and TG in intracytoplasmic micro/macrovesicles triggered by multiple factors especially in the presence of FA (Gómez-Lechón, Jover and Donato, 2009). Presence of lipid droplets and increased TG accumulation has been previously reported for *in vitro* cultures in the presence of long-term stimuli with (unsaturated) oleic acid (Antherieu *et al.*, 2011; Rogue *et al.*, 2014) and short-term treatment with a mixture of oleate and palmitate (O/P) (Dave, Tilles and Vemula, 2018; Chavez-Tapia, Rosso and Tiribelli, 2012; Brown *et al.*, 2013). The investigations on hepatic lipid content in NAFLD and NASH patients reveals increased levels of O/P and a trend that shows a shift towards O/P with the disease progression (Araya *et al.*, 2004). Hence, it was physiologically relevant to treat the HepaRG cells with O/P to stimulate steatosis. Moreover, we used a 2:1 ratio of O/P to mitigate the lipotoxic effects of palmitate treatment (Montell *et al.*, 2001; Wei, Wang and Pagliassotti, 2007). Here we report a 6-fold increase in TG content as well as increased lipid-filled vesicles in HepaRG cells after a 72h stimulus with 0.5mM of O/P (2:1). This *in vitro* model also showed a positive correlation with several genes associated with *de-novo* lipogenesis (Table 3.3) and inflammation. Thereby, providing a molecular signature resembling liver steatosis.

HepaRG cells treated with 0.5 mM O/P (2:1) showed a significant downregulation of PXR, CAR, and PPARα as well as the target CYP mRNA. Our results is in agreement with previous reports showing downregulation of PPARα in hepatic steatosis (Pawlak, Lefebvre and Staels, 2015). However, in contrast to our observations, Tanner et al. reported no change in mRNA of DMETs as well as the nuclear receptors such as PXR, CAR, and PPARα upon treatment with O/P (Tanner *et al.*, 2018). However, the FA treatment method was different from our experiment including the proportion of the two FA (1:1 O/P vs. 2:1 O/P), incubation period (24h vs. 72h) and concentration of FA treatment (125 μM vs. 500 μM), which might explain the discrepancies between the two studies.

Previous studies have reported the impact of hepatic steatosis on P450 enzymes and cell membrane transporters (Lake *et al.*, 2011; Fisher *et al.*, 2009; Hardwick *et al.*, 2013). In our NAFLD model, the gene expression of CYP including CYP3A4, 1A2, 2B6, 2C8, 2C9, 2C19 and 4A11 were significantly reduced. Additionally, we also observed a lower trend in CYP2A6, CYP2C8, CYP2C9 and CYP3A5 protein expression, however this was not statistically significant findings. Similar changes in gene expression were previously observed by Donato et al. (Donato *et al.*, 2006) using plated human hepatocytes, but these were only exposed to FFA for 14 h and the impact on other ADME proteins were not extensively explored. Kostrzewski et al. (Kostrzewski *et al.*, 2017), reported using a 3D human hepatocytes model of NAFLD that the metabolic rates of CYP3A4 and CYP2C9 to be reduced, with no changes in the genomic expression. These contrasted with our findings that show significant mRNA change, as our model is specific to early stages of steatosis that reverts during disease progression. In summary, utilizing an *in vitro* model of liver steatosis, our data demonstrate broad alterations in ADME, upon FA loading of HepaRG cells.

Increased infiltration of the FA in the liver results in an alteration in β-oxidation via TCA cycle and increased TG synthesis to compensate with the overload, that eventually leads to hepatocyte enlargement (Diraison, Moulin and Beylot, 2003). This transition is reflected in HepaRG FA treatment with the upregulation of aldo–keto reductase (AKR) 1B10 which regulates acetyl-CoA carboxylase (ACC)-catalyzed reaction step of de novo lipogenesis (Bitter *et al.*, 2015). Upregulation of PLIN2 is also noted, which coats lipid droplets and prevents TG to undergo lipolysis via LPL. Downregulation of PPARα which is important in lipid metabolism and regulates the expression of PLIN2 is noteworthy and consistent with the literature (McManaman *et al.*, 2013).

High-sensitivity C-reactive protein is a clinical predictor of NAFLD (Fierbinteanu-Braticevici *et al.*, 2011); in the current study, this marker was upregulated along with a minimal increase in serum amyloid A2 and IL-6 (Table 3.4). However, induction of acute-phase proteins upon FA treatment shows the early signs of progression from steatosis to the development of NASH. Despite the induction of a few inflammatory markers, there is no evidence to show the presence of oxidative stress in the cells, and hence this model is more representative of hepatic steatosis and co-culture with immune cells is warranted to optimize a model for NASH.

The nuclear receptor, Hepatocyte Nuclear Factor 4-alpha (HNF4α), regulated by the presence of FA is reported to modulate both xenobiotic and FA metabolism (Merrell and Cherrington, 2011). It is to be noted that, CYP2A6, 2B6, 2C8, 2C9, 2C19, 2D6 and 3A4/5 are controlled by upstream regulator HNF4α via PXR and CAR. Their xeno-sensing mechanism also plays a central role in lipid homeostasis and hepatocyte differentiation (Jover, Moya and Gómez-Lechón, 2009). HNF4α deficient mice are shown to have increased lipid accumulation in the liver and reduced serum cholesterol and TG levels (Hayhurst *et al.*, 2001). Using our transcriptomics data, we observed a downregulation of

HNF4α which may contribute to the progression to steatosis as well as dysregulation in DMET. However, most of the nuclear receptors and transcriptional regulator could not be identified in the proteomics assay due to the low abundance in cells.

Comparative analysis of transcriptomics and proteomics reveal important pattern of many proteins as seen in AKR1B10. However, there is a mismatch in DME mRNA and protein expression, this could be partly explained due to changes in posttranslational modifications and cellular location of proteins (Schwanhäusser *et al.*, 2011). Recently, Liu et al. reviewed the relationship between mRNA and proteins under different scenarios (Liu, Beyer and Aebersold, 2016). They concluded that the correlations may depend on many factors including applied techniques used in quantification of mRNA and protein, cell state, and posttranscriptional and posttranslational processes. These might have caused the observed variability. This may cause the change in expression pattern of the endpoints, primarily due to the time of harvest from the vitro culture. Also, the samples for the protein expression study in HepaRG cells were harvested from P17-P20 in comparison to P16 for mRNA samples. We believe the different passages are causing variability and leading to loss of significance in the assessment of CYP450 proteins that are generally present in low abundance in HepaRG cells.

In conclusion, our study presents an *in-vitro* model of steatosis demonstrating that lipid-loaded HepaRG cells are a suitable model for studying the effect of NAFLD on drug-metabolizing enzymes and transporters. Utilizing transcriptomics and proteomics data, we established the experimental model (2:1 O/P in HepaRG) showed a molecular signature that closely resembled hepatic steatosis leading to steatohepatitis. Additionally, we reported a significant downregulation of various DMET genes, along with a proposed mechanism for the possible initial 'hits' that leads to steatosis. FA-loaded HepaRG cells

can be a valuable *in-vitro* tool to model the disposition of new chemical entities in the disease state and may further aid in the appropriate design of new drugs for NAFLD.

5. Figures and Tables

Graphical Abstract

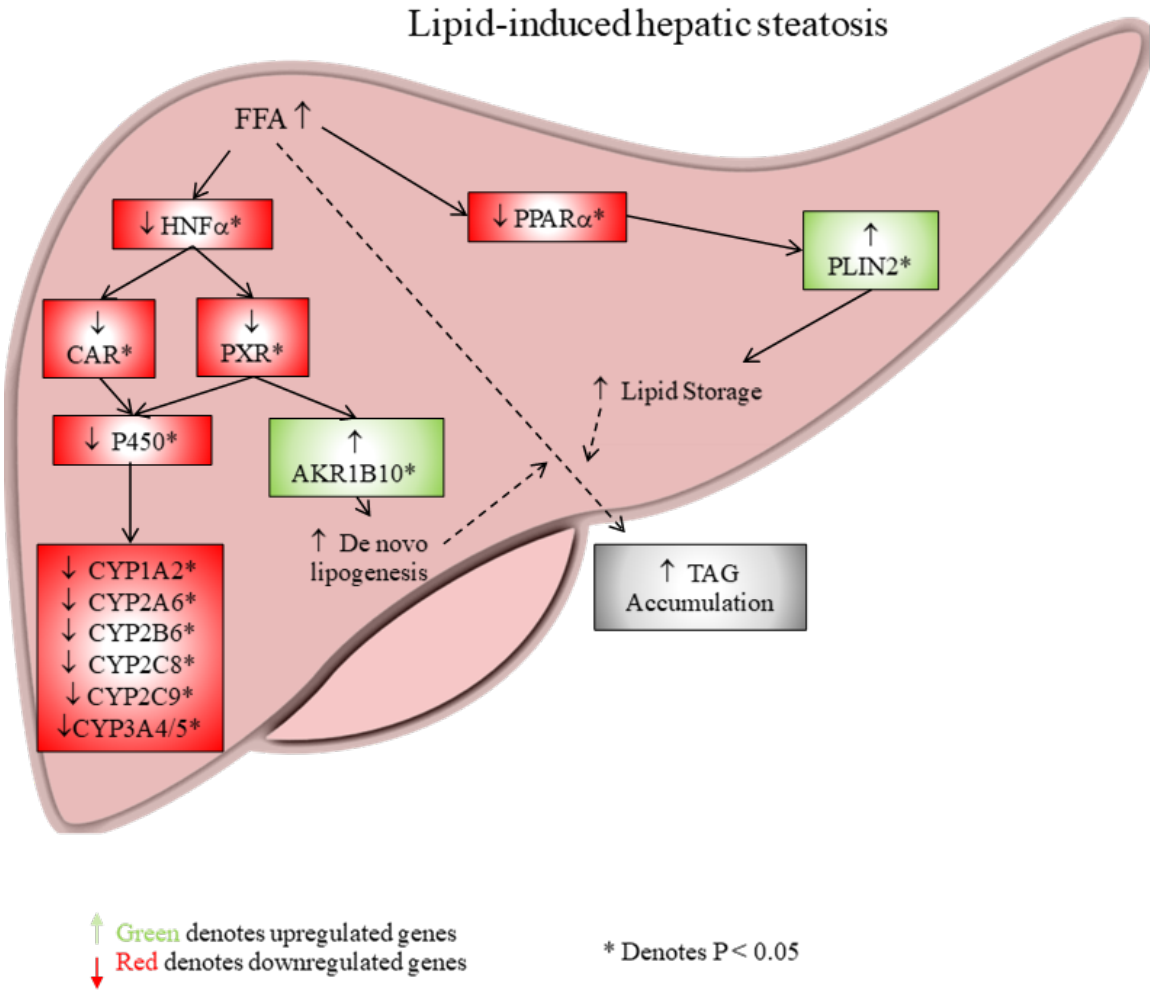

Figure 3.1. Model development of cellular steatosis by Oil Red O staining and cell viability assay.

(Fig 3.1-I) HepaRG cells were treated for 24h, 48h, 72h with control (A, C, E) vs. 0.5 µM of O/P 1:2 treatment (B, D, F). Images were taken under a light microscope at 20X magnification. The scale bars denote 85 µm. (Fig 1-II) For quantitative analysis of fat incorporation, Oil Red O stain was extracted and measured, OD @ 492 nm, following O/P treatment for 24, 48 and 72h respectively. Mean + SD (assayed in triplicate) are represented as relative to control cells. (Fig 1-III) Cells were incubated in the presence of increasing concentrations from 0.1 - 2.0 mM of O/P (1:2) mixture for 72h. Data from the WST-1 tests are depicted as means + SD (three biological replicates) and given relative to control cells (NC). PC (positive control) cells were treated with 0.1 mM of clotrimazole. A non-toxic and steatosis-inducing concentration of 0.5 mM O/P at 72h incubation was selected for further analyses

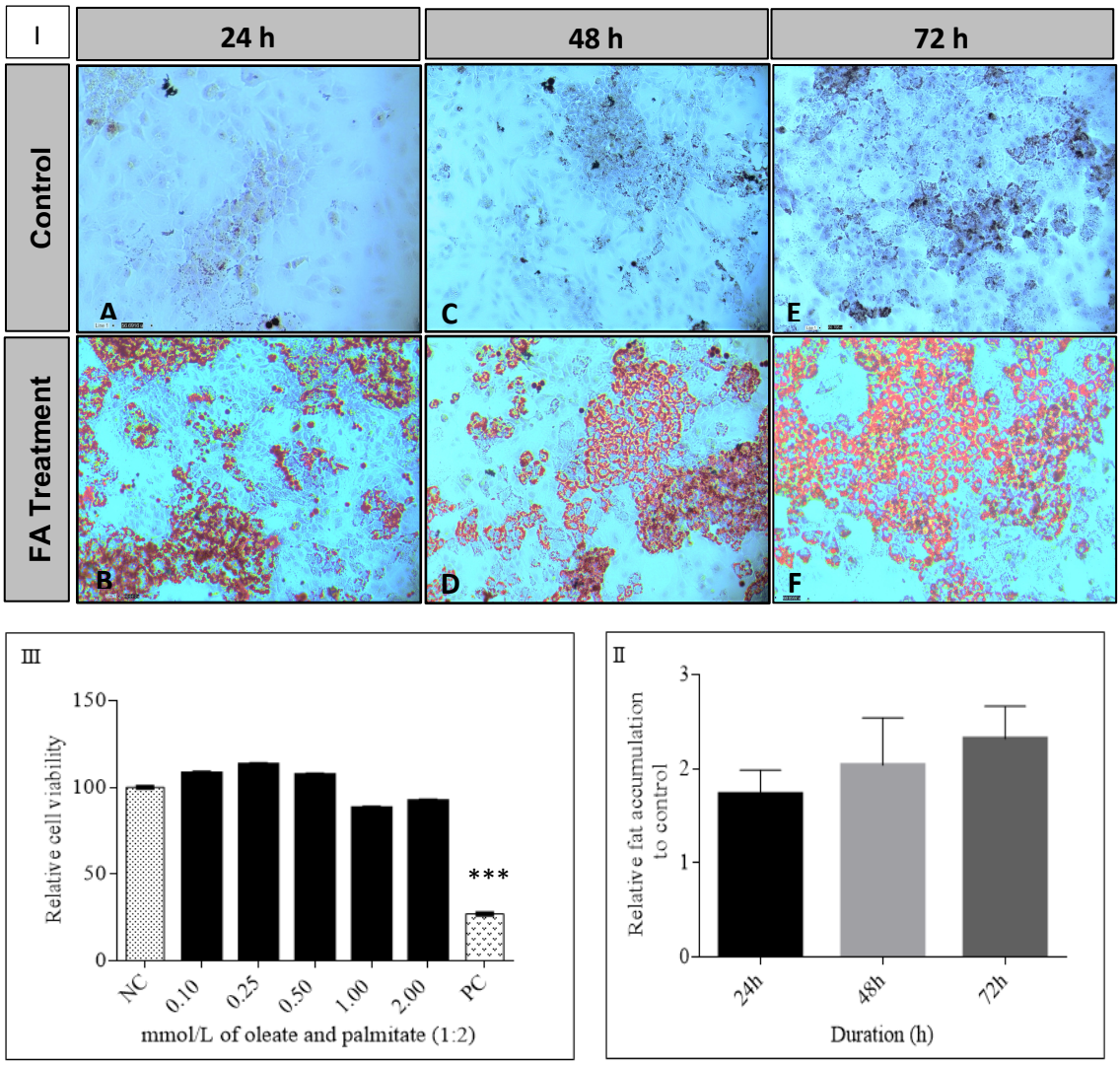

Figure 3.2: Model development of cellular steatosis by Oil Red O staining and cell viability assay.

(Fig 1-I) HepaRG cells were treated for 24h, 48h, 72h with control (A, C, E) vs. 0.5 µM of O/P 1:2 treatment (B, D, F). Images were taken under a light microscope at 20X magnification. The scale bars denote 85 µm. (Fig 1-II) For quantitative analysis of fat incorporation, Oil Red O stain was extracted and measured, OD @ 492 nm, following O/P treatment for 24, 48 and 72h respectively. Mean + SD (assayed in triplicate) are represented as relative to control cells. (Fig 1-III) Cells were incubated in the presence of increasing concentrations from 0.1 - 2.0 mM of O/P (1:2) mixture for 72h. Data from the WST-1 tests are depicted as means + SD (three biological replicates) and given relative to control cells (NC). PC (positive control) cells were treated with 0.1 mM of clotrimazole. A non-toxic and steatosis-inducing concentration of 0.5 mM O/P at 72h incubation was selected for further analyses (***$p<0.001$).

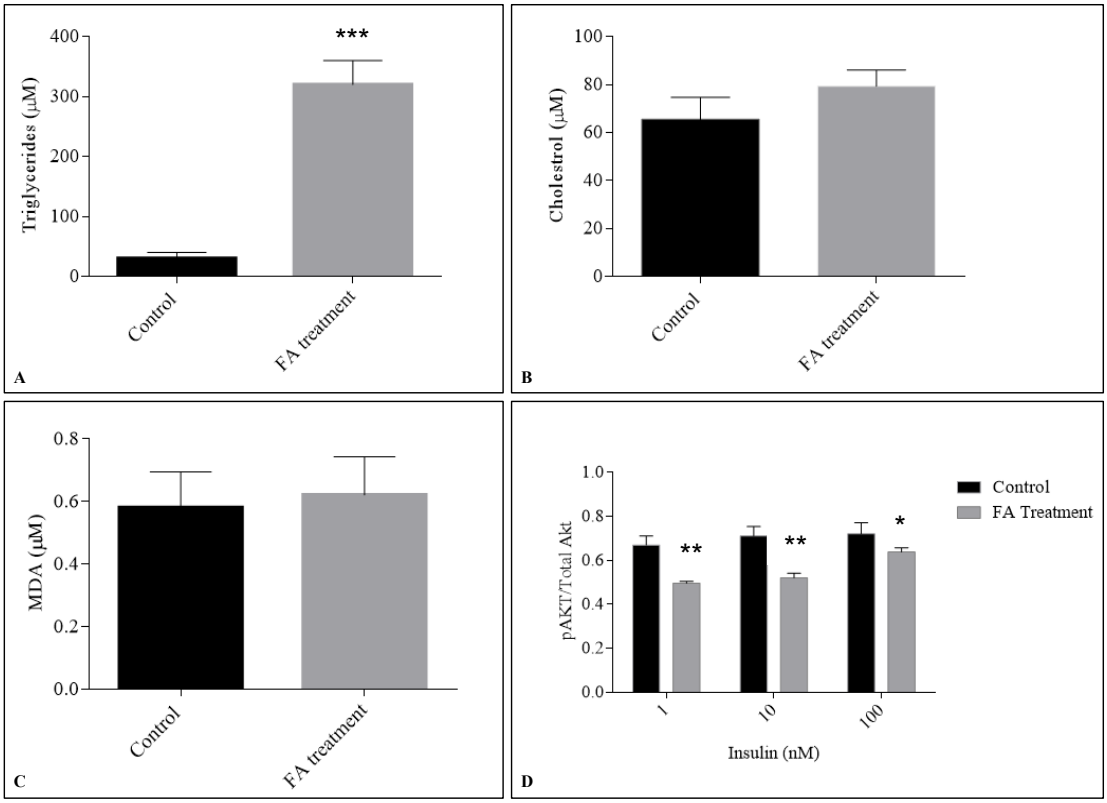

Figure 3.3: Model validation of cellular steatosis.

Differentiated HepaRG cells were treated with BSA alone (control) or with 0.5 μM palmitate: oleate (1: 2) (FA treatment) for 72 h prior to assaying triglyceride **(A)** and cholesterol **(B)** content. **(C)** The mean level of reactive oxygen species (ROS) detected via measuring malondialdehyde (MDA) as an indicator of oxidative stress is shown. **(D)** Cells were treated for 12h in serum free media (Starvation). Cells were then stimulated with insulin (1, 10, 100 nM) for 10 min before harvest. Phosphorylated Akt (Ser473) and total Akt were assayed in ELISA kit. Error bars represent standard deviation. All samples were assayed in triplicates. Asterisks indicate a significant (*P < 0.05; **P <0.01; ***P <0.001) difference compared to control.

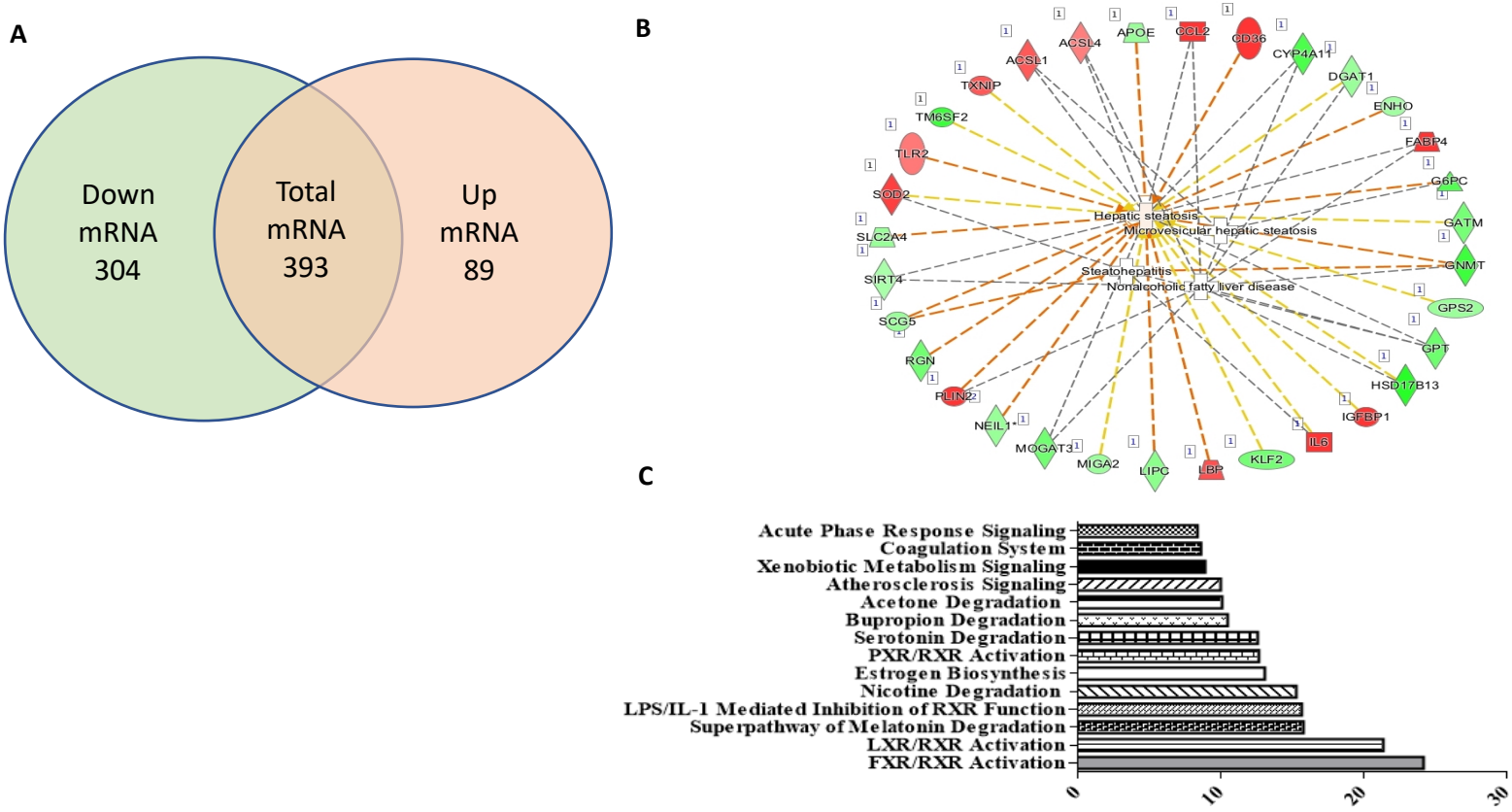

Figure 3.4: Differential expressed genes and proteins in HepaRG cells with FA treatment.

(A) Venn diagram indicating the differentially up- and down-regulated genes quantified using proteomics and transcriptomics studies; based on log FC > 0.58; pvalue < 0.05. (B) Venn diagram showing the differentially up- and down-regulated proteomics quantified using SWATH-MS. (C) Listed are the significant gene expression change that predict the model disease state to closely resemble steatosis, microvesicular hepatic steatosis, steatohepatitis as well as non-alcoholic fatty liver upon fatty acids addition to HepaRG cells. Red and green denote downregulated and upregulated genes, respectively. (D) The top 10 dysregulated pathway predicted by IPA.

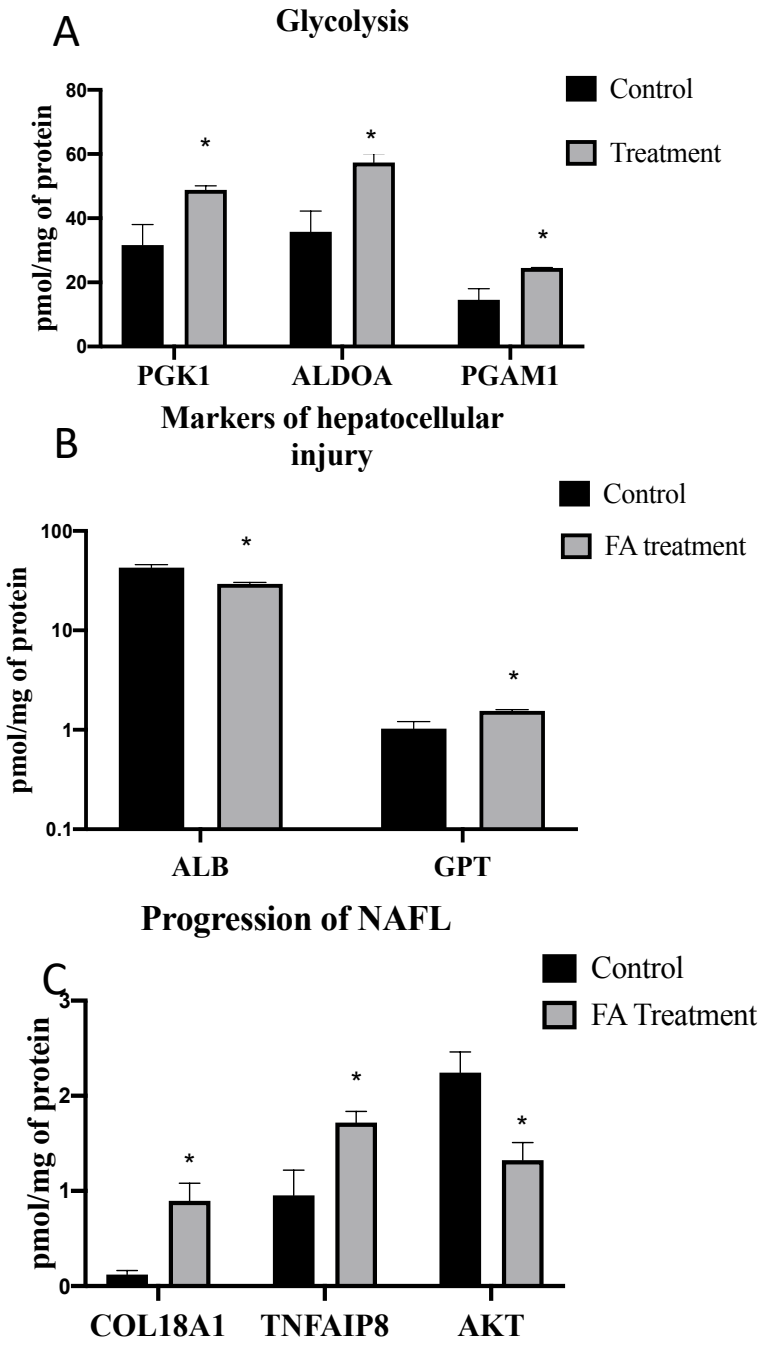

Figure 3.5: Changes in protein expression involved in glycolysis and in the progression of NAFL.

(A) The protein expression quantified by SWATH-MS of glycolytic proteins in control and FA treated HepaRG cells are shown. (B, C) The proteins involved in hepatocellular injury as well as markers for the progression of NAFL are shown. The expression of proteins is expressed as pmol/mg of protein. The figure is the mean of protein expression among the replicates; * denotes $p<0.05$.

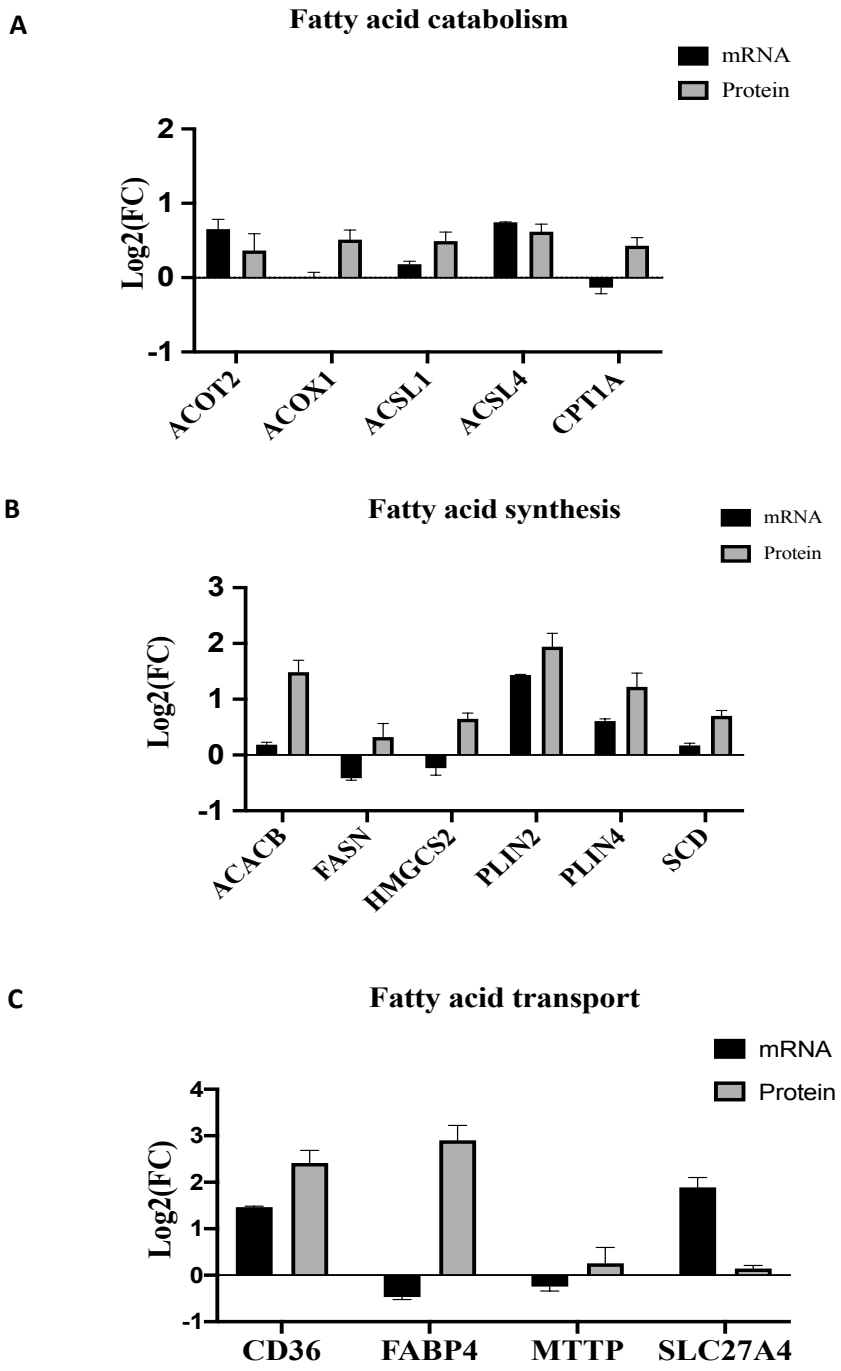

Figure 3.6: Transcripts and proteins involved in lipid catabolism, lipid synthesis and lipid transport.

The figure denotes the log fold change (FC) of expression in FA treatment vs. control. The mRNA and protein response are denoted as shown. The figure is shown as mean of log (FC) with SEM among the replicates.

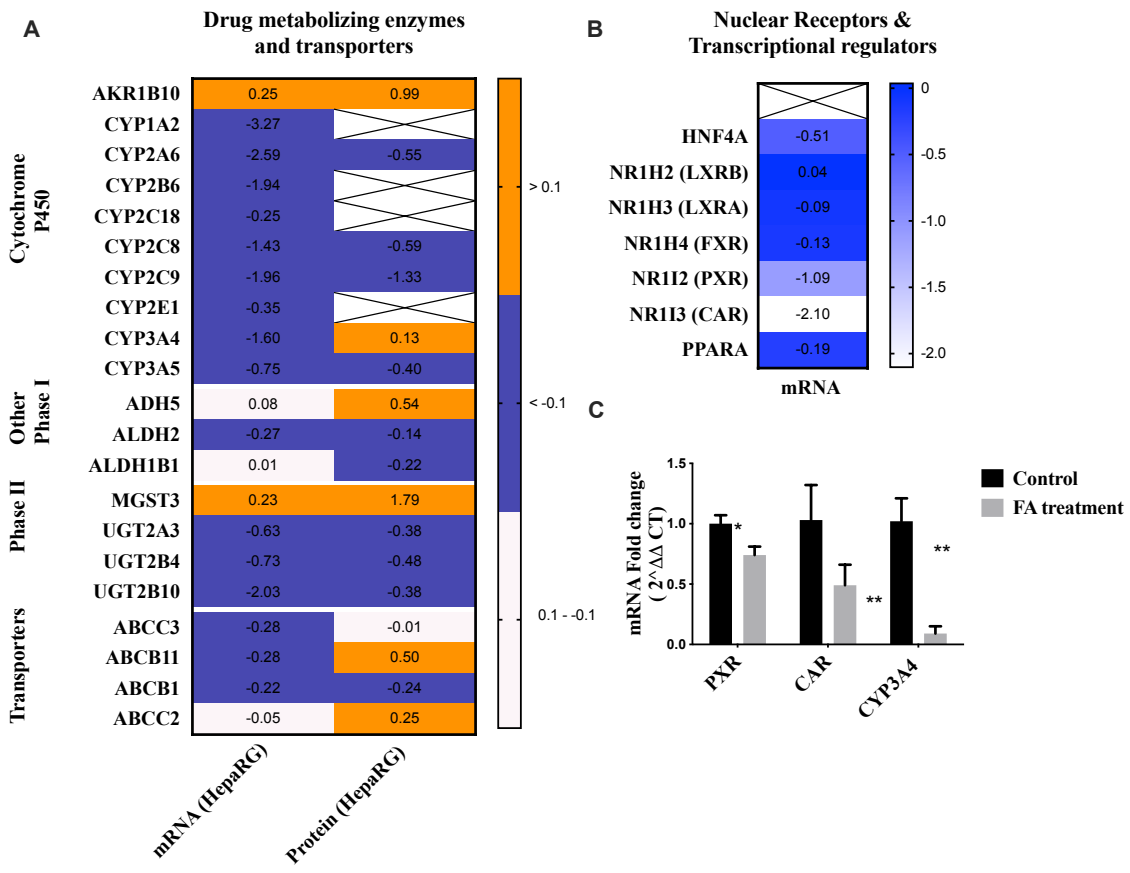

Figure 3.7: Fold change (logFC) of drug metabolizing genes and proteins upon treatment.

Differentiated HepaRG cells were treated with BSA alone (control) or 0.5 µM palmitate:oleate (1: 2) (FA treatment) for 72 h prior to SWATH-MS based proteomics analysis. **(A)** Heat map illustrating cytochrome P450 (CYP), Other Phase 1, Phase 2 and transporters mRNA and proteins are shown. Log FC of mRNA and protein expression as shown in the figure **(B)** Heat map of nuclear receptors and transcriptional regulators are expressed as log FC of mRNA and proteins. **(C)** mRNA expression of control and fatty acid treated HepaRG cells from q-PCR are expressed. All samples were assayed in triplicates and samples are represented as an average of log_2FC. X denotes absence of expression due to presence of proteins below the limit of quantification; * denotes $p<0.05$.

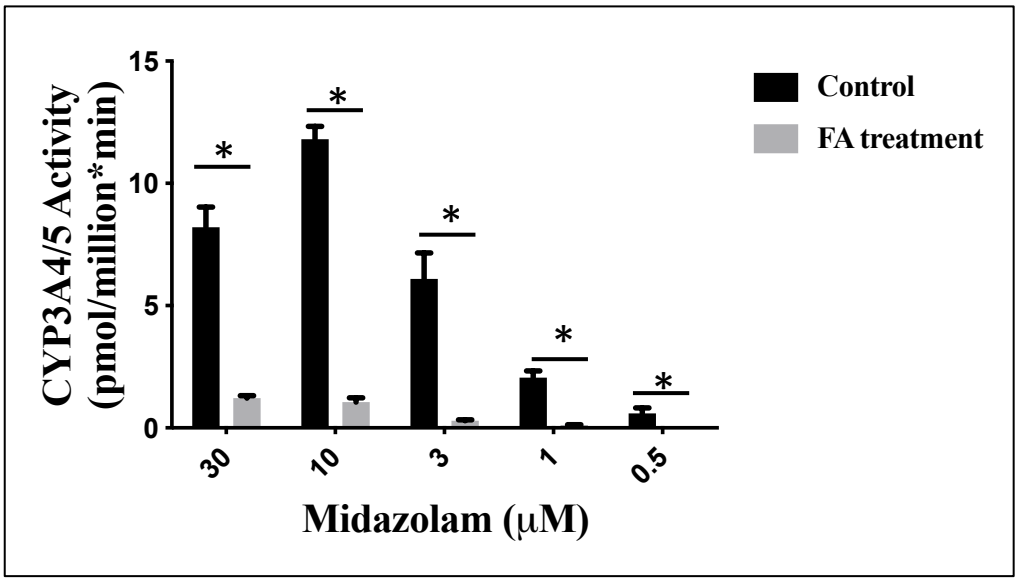

Figure 3.8: CYP3A4 functional activity in control and fatty acid treatment.

CYP3A4 enzyme activity assessed at different substrate concentration (30, 10, 3, 1, 0.5 μM) show significant downregulation upon FA treatment. The significance was calculated with $P < 0.05$.

Table 3.1: Comparison of transcriptional expression profile of lipid loaded HepaRG model and biobank liver samples

Gene Name	Description	RNA-seq/ FC	P value	q-PCR/ FC (Steatosis/cntrl)	P value
		(From current study)		*(Starmann et al., 2012)*	
SUSD2	sushi domain containing 2	6.53E-01	2.50E-16	1.47E+00	6.90E-02
ACSL4	acyl-CoA synthetase long-chain family member 4	1.51E+00	4.66E-11	2.11E+00	2.90E-02
ATF3	crystallin lambda 1	6.75E-01	2.18E-09	1.57E+00	1.74E-01
INMT	indolethylamine N-methyltransferase	3.24E-01	3.75E-09	1.03E+00	9.19E-01
MAT1A	methionine adenosyltransferase 1A	6.96E-01	6.23E-09	8.55E-01	3.21E-01
LUM	lumican	1.33E+00	1.83E-07	1.45E+00	6.10E-02
ACSL5	acyl-CoA synthetase long-chain family member 5	1.25E+00	4.28E-06	1.38E+00	4.90E-02
PNPLA3	patatin like phospholipase domain containing 3	7.62E-01	4.20E-05	2.15E+00	2.00E-03
SH3KBP1	SH3 domain containing kinase binding protein 1	1.31E+00	2.18E-04	1.38E+00	1.43E-01
ID3	inhibitor of DNA binding 3, HLH protein	7.10E-01	2.66E-04	1.01E+00	9.68E-01
SQSTM1	sequestosome 1	1.17E+00	2.30E-03	1.32E+00	1.30E-01
AKR1B10	aldo-keto reductase family 1 member B10	1.15E+00	6.89E-03	2.73E+00	7.00E-02

DGAT2	diacylglycerol O-acyltransferase 2	8.72E-01	1.49E-02	1.31E+00	2.31E-01
CYP3A43	cytochrome P450 family 3 subfamily A member 43	3.26E-01	3.41E-02	7.77E-01	4.05E-01
ITGA3	integrin subunit alpha 3	1.09E+00	1.01E-01	1.45E+00	1.33E-01
STMN1	stathmin 1	1.07E+00	2.22E-01	NA	NA
GSTM1	glutathione S-transferase mu 1	8.53E-01	2.27E-01	NA	NA
KRT18	keratin 18	1.05E+00	3.46E-01	1.18E+00	3.06E-01
CAT	catalase	9.58E-01	3.64E-01	7.22E-01	5.21E-01
NR0B2	nuclear receptor subfamily 0 group B member 2	6.44E-01	3.78E-01	2.27E+00	1.00E-03
DEFB1	defensin beta 1	9.46E-01	5.26E-01	2.36E+00	1.00E-03
LRRFIP2	LRR binding FLII interacting protein 2	1.01E+00	8.97E-01	1.80E+00	5.00E-03
KRT8	keratin 8	1.01E+00	9.07E-01	1.65E+00	4.10E-02
MAGED2	MAGE family member D2	1.00E+00	9.98E-01	1.46E+00	1.00E-02

Table 3. 2: Top 20 differentially regulated canonical pathway upon FA treatment

Ingenuity Canonical Pathways	-log(p-value)	Ratio	Molecules
FXR/RXR Activation	24.2	0.278	APOA4, TF, SLC10A1, SLC22A7, G6PC, RBP4, NR5A2, APOC1, APOC3, APOM, KNG1, BAAT, AMBP, APOL1, ALB, UGT2B4, LIPC, PKLR, ABCG5, FGA, ABCG8, CYP8B1, VTN, SAA2, MLXIPL, APOA2, FETUB, LCAT, APOE, SULT2A1, AHSG, HPR, CYP7A1, TTR, NR1I2
LXR/RXR Activation	21.4	0.264	IL1RL1, APOA4, TF, LBP, RBP4, IL6, APOC1, APOC3, APOM, KNG1, AMBP, APOL1, ALB, MYLIP, CCL7, TNFRSF1B, ABCG5, FGA, ABCG8, CD36, VTN, TLR3, SAA2, CCL2, MLXIPL, APOA2, LCAT, APOE, AHSG, HPR, CYP7A1, TTR
Superpathway of Melatonin Degradation	15.8	0.309	UGT2B4, CYP3A7, CYP4X1, UGT2B10, UGT1A7 (includes others), UGT2B7, CYP1A1, SULT1A2, CYP3A4, SULT1E1, UGT1A1, CYP2C9, CYP4B1, CYP2B6, CYP1A2, UGT2B15, UGT2A3, CYP2C8, CYP3A5, SULT2A1, CYP2A6 (includes others)
LPS/IL-1 Mediated Inhibition of RXR Function	15.7	0.158	IL1RL1, CYP3A7, CYP4A11, SLC10A1, SULT1A2, GSTA2, CYP3A4, LBP, NR5A2, APOC1, CYP2C9, FMO5, ACSL1, CYP2B6, GSTA1, CYP2A6 (includes others), GSTT2/GSTT2B, ALDH1L1, FABP4, LIPC, TNFRSF1B, ABCG5, ALDH6A1, SULT1E1, ABCG8, CHST11, ACSL4, CYP2C8, ACSL3, CYP3A5, APOE, SULT2A1, CYP7A1, CHST15, NR1I2
Nicotine Degradation	15.3	0.317	UGT2B4, CYP3A7, CYP4X1, UGT2B10, UGT1A7 (includes others), UGT2B7, CYP1A1, CYP3A4, UGT1A1, CYP2C9, FMO5, CYP4B1, CYP2B6, CYP1A2, UGT2B15, UGT2A3, CYP2C8, CYP3A5, INMT, CYP2A6 (includes others)
Estrogen Biosynthesis	13.1	0.385	HSD17B3, HSD17B2, CYP3A7, CYP4X1, CYP1A1, CYP3A4, CYP2C9, AKR1C4, CYP4B1, CYP2B6, CYP1A2, CYP2C8, CYP3A5, HSD17B13, CYP2A6 (includes others)
PXR/RXR Activation	12.7	0.277	CYP3A7, UGT1A7 (includes others), G6PC, GSTA2, CYP3A4, IL6, UGT1A1, CYP2C9, IGFBP1, CYP2B6, CYP1A2, CYP2C8, CYP3A5, SULT2A1, GSTA1, CYP2A6 (includes others), CYP7A1, NR1I2
Serotonin Degradation	12.6	0.253	UGT2B4, DHRS9, ALDH1L1, ADH1C, UGT2B10, UGT1A7 (includes others), UGT2B7, SULT1A2, ADH1A, ADH4, SULT1E1, UGT1A1, ADHFE1, ADH1B, UGT2B15, UGT2A3, PECR, ADH6, SULT2A1

Bupropion Degradation	10.5	0.44	CYP4B1, CYP2B6, CYP1A2, CYP3A7, CYP4X1, CYP2C8, CYP1A1, CYP3A5, CYP3A4, CYP2C9, CYP2A6 (includes others)
Acetone Degradation	10.1	0.407	CYP4B1, CYP2B6, CYP1A2, CYP3A7, CYP4X1, CYP2C8, CYP1A1, CYP3A5, CYP3A4, CYP2C9, CYP2A6 (includes others)
Atherosclerosis Signaling	10	0.165	APOA4, COL5A3, COL1A1, PLA2G6, RBP4, IL6, APOC1, APOC3, COL18A1, APOM, CXCL8, CD36, CCL2, APOA2, RARRES3, LCAT, PLA2G12B, APOE, ICAM1, APOL1, ALB
Xenobiotic Metabolism Signaling	8.92	0.105	UGT2B11, CYP3A7, UGT1A7 (includes others), CYP1A1, SULT1A2, GSTA2, CYP3A4, IL6, CYP2C9, FMO5, CYP2B6, GSTA1, FGFR3, UGT2B4, GSTT2/GSTT2B, ALDH1L1, UGT2B10, UGT2B7, ALDH6A1, MAP3K12, SULT1E1, UGT1A1, CHST11, CYP1A2, UGT2B15, CYP2C8, CYP3A5, SULT2A1, CHST15, NR1I2
Coagulation System	8.65	0.314	F12, F7, KNG1, FGB, FGA, F10, SERPINC1, PLG, SERPIND1, FGG, F13B
Acute Phase Response Signaling	8.4	0.129	TF, TNFRSF1B, SOCS3, FGA, LBP, RBP4, IL6, CRP, PLG, FGG, SOD2, HRG, SAA2, FGB, APOA2, SOCS1, AHSG, AMBP, RBP5, SERPIND1, ALB, TTR
Extrinsic Prothrombin Activation Pathway	8.33	0.5	F12, F7, FGB, FGA, F10, SERPINC1, FGG, F13B
Thyroid Hormone Metabolism II (via Conjugation and/or Degradation)	7.84	0.268	UGT2B4, UGT2B10, UGT1A7 (includes others), UGT2B7, UGT2B15, UGT2A3, SULT1A2, DIO1, SULT1E1, SULT2A1, UGT1A1
Intrinsic Prothrombin Activation Pathway	7.72	0.262	F12, KNG1, FGB, FGA, F10, COL5A3, COL1A1, SERPINC1, COL18A1, FGG, F13B
Bile Acid Biosynthesis, Neutral Pathway	7.64	0.538	AKR1D1, BAAT, CYP3A4, SCP2, AKR1C4, CYP8B1, CYP7A1
Ethanol Degradation II	7.21	0.27	DHRS9, ACSL1, ALDH1L1, ADH1C, ADH1B, ADH1A, PECR, ADH4, ADH6, ADHFE1
Hepatic Fibrosis / Hepatic Stellate Cell Activation	6.47	0.109	IL1RL1, COL7A1, BAMBI, KDR, TNFRSF1B, COL6A1, COL27A1, CTGF, COL5A3, IGFBP5, COL4A4, LBP, COL1A1, IL6, COL18A1, CXCL8, CCL2, MYL5, CXCL3, ICAM1

Table 3.3: Classification of differentially expressed genes upon FA treatment vs. control predictive of physiological disease state

Ensembl	Classification	Symbol	Entrez Gene Name	Expr Log Ratio	Expr p-value	Location	Family
ENSG00000151726	Hepatic Steatosis	ACSL1	acyl-CoA synthetase long chain family member 1	0.775	2.73E-08	Cytoplasm	enzyme
ENSG00000068366		ACSL4	acyl-CoA synthetase long chain family member 4	0.602	6.25E-05	Cytoplasm	enzyme
ENSG00000130203		APOE	apolipoprotein E	-0.723	9.44E-06	Extracellular Space	transporter
ENSG00000108691		CCL2	C-C motif chemokine ligand 2	1.48	1.08E-06	Extracellular Space	cytokine
ENSG00000135218		CD36	CD36 molecule	1.38	6.01E-05	Plasma Membrane	transmembrane receptor
ENSG00000187048		CYP4A11	cytochrome P450 family 4 subfamily A member 11	-1.28	2.48E-06	Cytoplasm	enzyme
ENSG00000185000		DGAT1	diacylglycerol O-acyltransferase 1	-0.711	4.67E-05	Cytoplasm	enzyme
ENSG00000170323		FABP4	fatty acid binding protein 4	1.86	2.98E-07	Cytoplasm	transporter

ENSG00000131482		G6PC	glucose-6-phosphatase catalytic subunit	-1.22	4.68E-07	Cytoplasm	phosphatase
ENSG00000171766		GATM	glycine amidinotransferase	-0.962	4.84E-08	Cytoplasm	enzyme
ENSG00000124713		GNMT	glycine N-methyltransferase	-1.39	0.00999	Cytoplasm	enzyme
ENSG00000170509		HSD17B13	hydroxysteroid 17-beta dehydrogenase 13	-1.54	2.35E-04	Extracellular Space	enzyme
ENSG00000146678		IGFBP1	insulin like growth factor binding protein 1	1.04	2.47E-05	Extracellular Space	other
ENSG00000136244		IL6	interleukin 6	1.28	0.00293	Extracellular Space	cytokine
ENSG00000127528		KLF2	Kruppel like factor 2	-0.98	0.0315	Nucleus	transcription regulator
ENSG00000129988		LBP	lipopolysaccharide binding protein	0.815	4.84E-08	Plasma Membrane	transporter
ENSG00000166035		LIPC	lipase C, hepatic type	-0.817	0.0059	Extracellular Space	enzyme
ENSG00000148343		MIGA2	mitoguardin 2	-0.716	0.00254	Plasma Membrane	other

ENSG00000106384	Microvesicular hepatic steatosis	MOGAT3	monoacylglycerol O-acyltransferase 3	-1.01	0.0303	Cytoplasm	enzyme
ENSG00000140398		NEIL1	Nei like DNA glycosylase 1	-0.731	0.024	Nucleus	enzyme
ENSG00000147872	Nonalcoholic fatty liver disease	PLIN2	perilipin 2	1.39	4.21E-10	Plasma Membrane	other
ENSG00000130988		RGN	regucalcin	-1.01	3.12E-04	Nucleus	enzyme
ENSG00000181856		SLC2A4	solute carrier family 2 member 4	-0.731	0.0184	Plasma Membrane	transporter
ENSG00000112096		SOD2	superoxide dismutase 2	0.896	7.55E-09	Cytoplasm	enzyme
ENSG00000137462	Steatohepatitis	TLR2	toll like receptor 2	0.628	0.00766	Plasma Membrane	transmembrane receptor
ENSG00000213996		TM6SF2	transmembrane 6 superfamily members 2	-1.41	0.00787	Plasma Membrane	other
ENSG00000265972		TXNIP	thioredoxin interacting protein	0.78	1.04E-06	Cytoplasm	other

Table 3.4: Classification of differentially expressed drug metabolizing genes and proteins upon FA treatment vs. control

Category	Feature	Description	Gene Name	logFC (mRNA)	p-value
	ENSG00000187758	alcohol dehydrogenase 1A (class I); alpha polypeptide	ADH1A	-1.30	1.01E-06
	ENSG00000196616	alcohol dehydrogenase 1B (class I); beta polypeptide	ADH1B	-1.45	5.93E-09
	ENSG00000248144	alcohol dehydrogenase 1C (class I); gamma polypeptide	ADH1C	-1.96	1.41E-08
	ENSG00000198099	alcohol dehydrogenase 4 (class II); pi polypeptide	ADH4	-1.70	6.97E-09
	ENSG00000140505	cytochrome P450 family 1 subfamily A member 2	CYP1A2	-3.25	1.25E-06
	ENSG00000255974	cytochrome P450 family 2 subfamily A member 6	CYP2A6	-2.61	2.49E-06
	ENSG00000197408	cytochrome P450 family 2 subfamily B member 6	CYP2B6	-2.02	1.60E-04
	ENSG00000138115	cytochrome P450 family 2 subfamily C member 8	CYP2C8	-1.48	1.62E-08
	ENSG00000138109	cytochrome P450 family 2 subfamily C member 9	CYP2C9	-2.00	2.17E-10
	ENSG00000160868	cytochrome P450 family 3 subfamily A member 4	CYP3A4	-1.87	1.60E-04
	ENSG00000106258	cytochrome P450 family 3 subfamily A member 5	CYP3A5	-0.80	4.29E-06
	ENSG00000160870	cytochrome P450 family 3 subfamily A member 7	CYP3A7	-1.29	2.98E-03
	ENSG00000187048	cytochrome P450 family 4 subfamily A member 11	CYP4A11	-1.28	2.48E-06
	ENSG00000186115	cytochrome P450 family 4 subfamily F member 2	CYP4F2	-1.34	2.73E-08

	ENSG00000167910	cytochrome P450 family 7 subfamily A member 1	CYP7A1	-2.71	7.50E-06
	ENSG00000180432	cytochrome P450 family 8 subfamily B member 1	CYP8B1	-1.61	6.80E-06
	ENSG00000131781	flavin containing monooxygenase 5	FMO5	-1.21	9.44E-06
	ENSG00000122787	aldo-keto reductase family 1-member B10	AKR1B10	2.17	1.41E-06
	ENSG00000198610	aldo-keto reductase family 1-member C	AKR1C4	-1.54	1.66E-07
Phase II	ENSG00000241119	UDP glucuronosyltransferase family 1-member A9	UGT1A9	-1.40	7.55E-07
	ENSG00000196620	UDP glucuronosyltransferase family 2-member B15	UGT2B15	-0.95	9.82E-08
	ENSG00000135220	UDP glucuronosyltransferase family 2-member A3	UGT2A3	-0.68	4.72E-06
	ENSG00000171234	UDP glucuronosyltransferase family 2-member B7	UGT2B7	-1.09	8.72E-08
	ENSG00000213759	UDP glucuronosyltransferase family 2-member B11	UGT2B11	-1.27	1.39E-04
	ENSG00000133433	glutathione S-transferase theta 2B	GSTT2B	-1.07	8.48E-03
	ENSG00000243955	glutathione S-transferase alpha 1	GSTA1	-0.91	6.51E-08
	ENSG00000105398	sulfotransferase family 2A member 1	SULT2A1	-0.75	6.91E-07
	ENSG00000109193	sulfotransferase family 1E member 1	SULT1E1	-1.41	1.64E-04
Transporters	ENSG00000121270	ATP binding cassette subfamily C member 11	ABCC11	-1.36	3.13E-05
	ENSG00000143921	ATP binding cassette subfamily G member 8	ABCG8	-0.97	6.30E-04
	ENSG00000138075	ATP binding cassette subfamily G member 5	ABCG5	-1.66	8.38E-04

	ENSG00000117834	solute carrier family 5 member 9	SLC5A9	-1.23	1.22E-05
	ENSG00000100652	solute carrier family 10 member 1	SLC10A1	-1.63	5.91E-10
	ENSG00000137491	solute carrier organic anion transporter family member 2B1	SLCO2B1	-1.49	1.02E-06
	ENSG00000134538	solute carrier organic anion transporter family member 1B1	SLCO1B1	-0.39	5.49E-02
Nuclear Receptors & Transcription factors	ENSG00000126368	nuclear receptor subfamily 1 group D member 1	NR1D1	-0.52	2.11E-03
	ENSG00000144852	nuclear receptor subfamily 1 group I member 2	NR1I2	-1.18	3.70E-06
	ENSG00000101076	hepatocyte nuclear factor 4 alpha	HNF4A	-0.55	7.03E-05

6. Reference

1 Aljomah, G., Baker, S. S., Liu, W., Kozielski, R., Oluwole, J., Lupu, B., Baker, R. D. and Zhu, L. (2015) 'Induction of CYP2E1 in non-alcoholic fatty liver diseases', *Exp Mol Pathol,* 99(3), pp. 677-81.

2 Aninat, C., Piton, A., Glaise, D., Le Charpentier, T., Langouet, S., Morel, F., Guguen-Guillouzo, C. and Guillouzo, A. (2006) 'Expression of cytochromes P450, conjugating enzymes and nuclear receptors in human hepatoma HepaRG cells', *Drug Metab Dispos,* 34(1), pp. 75-83.

3 Anthérieu, S., Rogue, A., Fromenty, B., Guillouzo, A. and Robin, M. A. (2011) 'Induction of vesicular steatosis by amiodarone and tetracycline is associated with up-regulation of lipogenic genes in HepaRG cells', *Hepatology,* 53(6), pp. 1895-905.

4 Araya, J., Rodrigo, R., Videla, L. A., Thielemann, L., Orellana, M., Pettinelli, P. and Poniachik, J. (2004) 'Increase in long-chain polyunsaturated fatty acid n - 6/n - 3 ratio in relation to hepatic steatosis in patients with non-alcoholic fatty liver disease', *Clin Sci (Lond),* 106(6), pp. 635-43.

5 Bellentani, S. (2017) 'The epidemiology of non-alcoholic fatty liver disease', *Liver Int,* 37 Suppl 1, pp. 81-84.

6 Bitter, A., Rümmele, P., Klein, K., Kandel, B. A., Rieger, J. K., Nüssler, A. K., Zanger, U. M., Trauner, M., Schwab, M. and Burk, O. (2015) 'Pregnane X receptor activation and silencing promote steatosis of human hepatic cells by distinct lipogenic mechanisms', *Arch Toxicol,* 89(11), pp. 2089-103.

7 Brodosi, L., Marchignoli, F., Petroni, M. L. and Marchesini, G. (2016) 'NASH: A glance at the landscape of pharmacological treatment', *Ann Hepatol,* 15(5), pp. 673-81.

8 Brown, M. V., Compton, S. A., Milburn, M. V., Lawton, K. A. and Cheatham, B. (2013) 'Metabolomic signatures in lipid-loaded HepaRGs reveal pathways involved in steatotic progression', *Obesity (Silver Spring),* 21(12), pp. E561-70.

9 Brunt, E. M. and Tiniakos, D. G. (2010) 'Histopathology of nonalcoholic fatty liver disease', *World J Gastroenterol,* 16(42), pp. 5286-96.

10 Bulutoglu, B., Rey-Bedón, C., Kang, Y. B. A., Mert, S., Yarmush, M. L. and Usta, O. B. (2019) 'A microfluidic patterned model of non-alcoholic fatty liver disease: applications to disease progression and zonation', *Lab Chip,* 19(18), pp. 3022-3031.

11 Caldwell, S. and Argo, C. (2010) 'The natural history of non-alcoholic fatty liver disease', *Dig Dis,* 28(1), pp. 162-8.

12 Chavez-Tapia, N. C., Rosso, N. and Tiribelli, C. (2012) 'Effect of intracellular lipid accumulation in a new model of non-alcoholic fatty liver disease', *BMC Gastroenterol,* 12, pp. 20.

13 Cobbina, E. and Akhlaghi, F. (2017) 'Non-alcoholic fatty liver disease (NAFLD) - pathogenesis, classification, and effect on drug metabolizing enzymes and transporters', *Drug Metab Rev,* 49(2), pp. 197-211.

14 Cohen, J. C., Horton, J. D. and Hobbs, H. H. (2011) 'Human fatty liver disease: old questions and new insights', *Science,* 332(6037), pp. 1519-23.

15 Cydylo, M. A., Davis, A. T. and Kavanagh, K. (2017) 'Fatty liver promotes fibrosis in monkeys consuming high fructose', *Obesity (Silver Spring),* 25(2), pp. 290-293.

16 Dave, T., Tilles, A. W. and Vemula, M. (2018) 'A Cell-Based Assay to Investigate Hypolipidemic Effects of Nonalcoholic Fatty Liver Disease Therapeutics', *SLAS Discov,* 23(3), pp. 274-282.

17 Diraison, F., Moulin, P. and Beylot, M. (2003) 'Contribution of hepatic de novo lipogenesis and reesterification of plasma non esterified fatty acids to plasma triglyceride synthesis during non-alcoholic fatty liver disease', *Diabetes Metab,* 29(5), pp. 478-85.

18 Donato, M. T., Lahoz, A., Jiménez, N., Pérez, G., Serralta, A., Mir, J., Castell, J. V. and Gómez-Lechón, M. J. (2006) 'Potential impact of steatosis on cytochrome P450 enzymes of human hepatocytes isolated from fatty liver grafts', *Drug Metab Dispos,* 34(9), pp. 1556-62.

19 Dostalek, M., Macwan, J. S., Chitnis, S. D., Ionita, I. A. and Akhlaghi, F. (2010) 'Development and validation of a rapid and sensitive assay for simultaneous quantification of midazolam, 1'-hydroxymidazolam, and 4-hydroxymidazolam by liquid chromatography coupled to tandem mass-spectrometry', *J Chromatogr B Analyt Technol Biomed Life Sci,* 878(19), pp. 1629-33.

20 Farrow, S. C. and Facchini, P. J. (2013) 'Dioxygenases catalyze O-demethylation and O,O-demethylenation with widespread roles in benzylisoquinoline alkaloid metabolism in opium poppy', *J Biol Chem,* 288(40), pp. 28997-9012.

21 Fierbinteanu-Braticevici, C., Baicus, C., Tribus, L. and Papacocea, R. (2011) 'Predictive factors for nonalcoholic steatohepatitis (NASH) in patients with nonalcoholic fatty liver disease (NAFLD)', *J Gastrointestin Liver Dis,* 20(2), pp. 153-9.

22 Fisher, C. D., Lickteig, A. J., Augustine, L. M., Ranger-Moore, J., Jackson, J. P., Ferguson, S. S. and Cherrington, N. J. (2009) 'Hepatic cytochrome P450 enzyme alterations in humans with progressive stages of nonalcoholic fatty liver disease', *Drug Metab Dispos,* 37(10), pp. 2087-94.

23 Gan, L., Chitturi, S. and Farrell, G. C. (2011) 'Mechanisms and implications of age-related changes in the liver: nonalcoholic Fatty liver disease in the elderly', *Curr Gerontol Geriatr Res,* 2011, pp. 831536.

24 Gerets, H. H., Tilmant, K., Gerin, B., Chanteux, H., Depelchin, B. O., Dhalluin, S. and Atienzar, F. A. (2012) 'Characterization of primary human hepatocytes, HepG2 cells, and HepaRG cells at the mRNA level and CYP activity in response to inducers and

their predictivity for the detection of human hepatotoxins', *Cell Biol Toxicol,* 28(2), pp. 69-87.

25 Ghazalpour, A., Bennett, B., Petyuk, V. A., Orozco, L., Hagopian, R., Mungrue, I. N., Farber, C. R., Sinsheimer, J., Kang, H. M., Furlotte, N., Park, C. C., Wen, P. Z., Brewer, H., Weitz, K., Camp, D. G., Pan, C., Yordanova, R., Neuhaus, I., Tilford, C., Siemers, N., Gargalovic, P., Eskin, E., Kirchgessner, T., Smith, D. J., Smith, R. D. and Lusis, A. J. (2011) 'Comparative analysis of proteome and transcriptome variation in mouse', *PLoS Genet,* 7(6), pp. e1001393.

26 Gorden, D. L., Myers, D. S., Ivanova, P. T., Fahy, E., Maurya, M. R., Gupta, S., Min, J., Spann, N. J., McDonald, J. G., Kelly, S. L., Duan, J., Sullards, M. C., Leiker, T. J., Barkley, R. M., Quehenberger, O., Armando, A. M., Milne, S. B., Mathews, T. P., Armstrong, M. D., Li, C., Melvin, W. V., Clements, R. H., Washington, M. K., Mendonsa, A. M., Witztum, J. L., Guan, Z., Glass, C. K., Murphy, R. C., Dennis, E. A., Merrill, A. H., Russell, D. W., Subramaniam, S. and Brown, H. A. (2015) 'Biomarkers of NAFLD progression: a lipidomics approach to an epidemic', *J Lipid Res,* 56(3), pp. 722-36.

27 Guillouzo, A., Corlu, A., Aninat, C., Glaise, D., Morel, F. and Guguen-Guillouzo, C. (2007) 'The human hepatoma HepaRG cells: a highly differentiated model for studies of liver metabolism and toxicity of xenobiotics', *Chem Biol Interact,* 168(1), pp. 66-73.

28 Gómez-Lechón, M. J., Jover, R. and Donato, M. T. (2009) 'Cytochrome p450 and steatosis', *Curr Drug Metab,* 10(7), pp. 692-9.

29 Hardwick, R. N., Ferreira, D. W., More, V. R., Lake, A. D., Lu, Z., Manautou, J. E., Slitt, A. L. and Cherrington, N. J. (2013) 'Altered UDP-glucuronosyltransferase and sulfotransferase expression and function during progressive stages of human nonalcoholic fatty liver disease', *Drug Metab Dispos,* 41(3), pp. 554-61.

30 Hayhurst, G. P., Lee, Y. H., Lambert, G., Ward, J. M. and Gonzalez, F. J. (2001) 'Hepatocyte nuclear factor 4alpha (nuclear receptor 2A1) is essential for maintenance of hepatic gene expression and lipid homeostasis', *Mol Cell Biol,* 21(4), pp. 1393-403.

31 Jamwal, R., Barlock, B. J., Adusumalli, S., Ogasawara, K., Simons, B. L. and Akhlaghi, F. (2017) 'Multiplex and Label-Free Relative Quantification Approach for Studying Protein Abundance of Drug Metabolizing Enzymes in Human Liver Microsomes Using SWATH-MS', *J Proteome Res,* 16(11), pp. 4134-4143.

32 Jamwal, R., de la Monte, S. M., Ogasawara, K., Adusumalli, S., Barlock, B. B. and Akhlaghi, F. (2018) 'Nonalcoholic Fatty Liver Disease and Diabetes Are Associated with Decreased CYP3A4 Protein Expression and Activity in Human Liver', *Mol Pharm,* 15(7), pp. 2621-2632.

33 Jover, R., Moya, M. and Gómez-Lechón, M. J. (2009) 'Transcriptional regulation of cytochrome p450 genes by the nuclear receptor hepatocyte nuclear factor 4-alpha', *Curr Drug Metab,* 10(5), pp. 508-19.

34 Koo, S. H. (2013) 'Nonalcoholic fatty liver disease: molecular mechanisms for the hepatic steatosis', *Clin Mol Hepatol,* 19(3), pp. 210-5.

35 Kostrzewski, T., Cornforth, T., Snow, S. A., Ouro-Gnao, L., Rowe, C., Large, E. M. and Hughes, D. J. (2017) 'Three-dimensional perfused human', *World J Gastroenterol,* 23(2), pp. 204-215.

36 Lake, A. D., Novak, P., Fisher, C. D., Jackson, J. P., Hardwick, R. N., Billheimer, D. D., Klimecki, W. T. and Cherrington, N. J. (2011) 'Analysis of global and absorption, distribution, metabolism, and elimination gene expression in the progressive stages of human nonalcoholic fatty liver disease', *Drug Metab Dispos,* 39(10), pp. 1954-60.

37 Larter, C. Z. and Yeh, M. M. (2008) 'Animal models of NASH: getting both pathology and metabolic context right', *J Gastroenterol Hepatol,* 23(11), pp. 1635-48.

38 Lecluyse, E. L. and Alexandre, E. (2010) 'Isolation and culture of primary hepatocytes from resected human liver tissue', *Methods Mol Biol,* 640, pp. 57-82.

39 Li, H., Toth, E. and Cherrington, N. J. (2018) 'Alcohol Metabolism in the Progression of Human Nonalcoholic Steatohepatitis', *Toxicol Sci,* 164(2), pp. 428-438.

40 Liu, Y., Beyer, A. and Aebersold, R. (2016) 'On the Dependency of Cellular Protein Levels on mRNA Abundance', *Cell,* 165(3), pp. 535-50.

41 Martignoni, M., Groothuis, G. M. and de Kanter, R. (2006) 'Species differences between mouse, rat, dog, monkey and human CYP-mediated drug metabolism, inhibition and induction', *Expert Opin Drug Metab Toxicol,* 2(6), pp. 875-94.

42 McManaman, J. L., Bales, E. S., Orlicky, D. J., Jackman, M., MacLean, P. S., Cain, S., Crunk, A. E., Mansur, A., Graham, C. E., Bowman, T. A. and Greenberg, A. S. (2013) 'Perilipin-2-null mice are protected against diet-induced obesity, adipose inflammation, and fatty liver disease', *J Lipid Res,* 54(5), pp. 1346-59.

43 Merrell, M. D. and Cherrington, N. J. (2011) 'Drug metabolism alterations in nonalcoholic fatty liver disease', *Drug Metab Rev,* 43(3), pp. 317-34.

44 Montell, E., Turini, M., Marotta, M., Roberts, M., Noé, V., Ciudad, C. J., Macé, K. and Gómez-Foix, A. M. (2001) 'DAG accumulation from saturated fatty acids desensitizes insulin stimulation of glucose uptake in muscle cells', *Am J Physiol Endocrinol Metab,* 280(2), pp. E229-37.

45 Moylan, C. A., Pang, H., Dellinger, A., Suzuki, A., Garrett, M. E., Guy, C. D., Murphy, S. K., Ashley-Koch, A. E., Choi, S. S., Michelotti, G. A., Hampton, D. D., Chen, Y., Tillmann, H. L., Hauser, M. A., Abdelmalek, M. F. and Diehl, A. M. (2014) 'Hepatic gene expression profiles differentiate presymptomatic patients with mild versus severe nonalcoholic fatty liver disease', *Hepatology,* 59(2), pp. 471-82.

46 Noureddin, M., Zhang, A. and Loomba, R. (2016) 'Promising therapies for treatment of nonalcoholic steatohepatitis', *Expert Opin Emerg Drugs,* 21(3), pp. 343-57.

47 Ohata, K., Hamasaki, K., Toriyama, K., Matsumoto, K., Saeki, A., Yanagi, K., Abiru, S., Nakagawa, Y., Shigeno, M., Miyazoe, S., Ichikawa, T., Ishikawa, H., Nakao, K. and Eguchi, K. (2003) 'Hepatic steatosis is a risk factor for hepatocellular carcinoma in patients with chronic hepatitis C virus infection', *Cancer,* 97(12), pp. 3036-43.

48 Parafati, M., Kirby, R. J., Khorasanizadeh, S., Rastinejad, F. and Malany, S. (2018) 'A nonalcoholic fatty liver disease model in human induced pluripotent stem cell-derived hepatocytes, created by endoplasmic reticulum stress-induced steatosis', *Dis Model Mech,* 11(9).

49 Pawlak, M., Lefebvre, P. and Staels, B. (2015) 'Molecular mechanism of PPARα action and its impact on lipid metabolism, inflammation and fibrosis in non-alcoholic fatty liver disease', *J Hepatol,* 62(3), pp. 720-33.

50 Prueksaritanont, T. (2010) 'Use of In Vivo Animal Models to Assess Drug–Drug Interactions; Progress and Future Challenges', in Pang, K.S.R., A. D.; Peter, R. M. (ed.) *Enzyme- and Transporter-Based Drug–Drug Interactions*. New York City: Springer.

51 Rogue, A., Anthérieu, S., Vluggens, A., Umbdenstock, T., Claude, N., de la Moureyre-Spire, C., Weaver, R. J. and Guillouzo, A. (2014) 'PPAR agonists reduce steatosis in oleic acid-overloaded HepaRG cells', *Toxicol Appl Pharmacol,* 276(1), pp. 73-81.

52 Rubin, K., Janefeldt, A., Andersson, L., Berke, Z., Grime, K. and Andersson, T. B. (2015) 'HepaRG cells as human-relevant in vitro model to study the effects of inflammatory stimuli on cytochrome P450 isoenzymes', *Drug Metab Dispos,* 43(1), pp. 119-25.

53 Saravanakumar, A., Sadighi, A., Ryu, R. and Akhlaghi, F. (2019) 'Physicochemical Properties, Biotransformation, and Transport Pathways of Established and Newly Approved Medications: A Systematic Review of the Top 200 Most Prescribed Drugs vs. the FDA-Approved Drugs Between 2005 and 2016', *Clin Pharmacokinet,* 58(10), pp. 1281-1294.

54 Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W. and Selbach, M. (2011) 'Global quantification of mammalian gene expression control', *Nature,* 473(7347), pp. 337-42.

55 Starmann, J., Fälth, M., Spindelböck, W., Lanz, K. L., Lackner, C., Zatloukal, K., Trauner, M. and Sültmann, H. (2012) 'Gene expression profiling unravels cancer-related hepatic molecular signatures in steatohepatitis but not in steatosis', *PLoS One,* 7(10), pp. e46584.

56 Stepanova, M., Hossain, N., Afendy, A., Perry, K., Goodman, Z. D., Baranova, A. and Younossi, Z. (2010) 'Hepatic gene expression of Caucasian and African-American patients with obesity-related non-alcoholic fatty liver disease', *Obes Surg,* 20(5), pp. 640-50.

57 Takahashi, Y., Soejima, Y. and Fukusato, T. (2012) 'Animal models of nonalcoholic fatty liver disease/nonalcoholic steatohepatitis', *World J Gastroenterol,* 18(19), pp. 2300-8.

58 Tanner, N., Kubik, L., Luckert, C., Thomas, M., Hofmann, U., Zanger, U. M., Bohmert, L., Lampen, A. and Braeuning, A. (2018) 'Regulation of drug metabolism by the interplay of inflammatory signaling, steatosis, and xeno-sensing receptors in HepaRG cells', *Drug Metab Dispos*.

59 Veteläinen, R., van Vliet, A., Gouma, D. J. and van Gulik, T. M. (2007) 'Steatosis as a risk factor in liver surgery', *Ann Surg,* 245(1), pp. 20-30.

60 Wei, Y., Wang, D. and Pagliassotti, M. J. (2007) 'Saturated fatty acid-mediated endoplasmic reticulum stress and apoptosis are augmented by trans-10, cis-12-conjugated linoleic acid in liver cells', *Mol Cell Biochem,* 303(1-2), pp. 105-13.

61 Younossi, Z. M., Gorreta, F., Ong, J. P., Schlauch, K., Del Giacco, L., Elariny, H., Van Meter, A., Younoszai, A., Goodman, Z., Baranova, A., Christensen, A., Grant, G. and Chandhoke, V. (2005) 'Hepatic gene expression in patients with obesity-related non-alcoholic steatohepatitis', *Liver Int,* 25(4), pp. 760-71.

62 Zhao, S., Xi, L., Quan, J., Xi, H., Zhang, Y., von Schack, D., Vincent, M. and Zhang, B. (2016) 'QuickRNASeq lifts large-scale RNA-seq data analyses to the next level of automation and interactive visualization', *BMC Genomics,* 17, pp. 39.

In Preparation for submission to Clinical Proteomics

MANUSCRIPT 4

Data-Independent Acquisition-Based Quantitative Proteomic Analysis Reveals Potential Biomarkers of Non-alcoholic fatty liver disease

Anitha Saravanakumar[1], Benjamin B Barlock[1], Rohitash Jamwal[1], Emily Martell[1], Angela Slitt[1], Fatemeh Akhlaghi[1]

[1]Biomedical and Pharmaceutical Sciences, College of Pharmacy, University of Rhode Island, Kingston, RI, 02881, USA

Corresponding author at:

Fatemeh Akhlaghi, Ph.D.

Clinical Pharmacokinetics Research Laboratory

Department of Biomedical and Pharmaceutical Sciences

The University of Rhode Island, 495A College of Pharmacy, 7 Greenhouse Road,

Kingston, RI 02881, United States.

Email address: fatemeh@uri.edu

Abbreviation

ANXA1 - Annexin A1

APOA4 - Apolipoprotein A-IV

APOE - Apolipoprotein E

APPL1 - DCC-interacting protein 13-alpha

ATP1A2 - Sodium/potassium-transporting ATPase subunit alpha-2

BCAT2- Branched-chain-amino-acid aminotransferase

CALD1- Caldesmon

CLTB - Clathrin light chain B

DBNL - Drebrin-like protein

EIF4H - Eukaryotic translation initiation factor 4H

EIF5B - Eukaryotic translation initiation factor 5B

G6PC - Glucose-6-phosphatase

GMPPA - Mannose-1-phosphate guanyltransferase alpha

GSTA4 - Glutathione S-transferase A4

H2AFY2 - Core histone macro-H2A.2

HMGB1 - High mobility group protein B1

KRT10 - Keratin, type I cytoskeletal 10

LTF - Lactotransferrin

NAFLD - Nonalcoholic fatty liver disease

NASH - Nonalcoholic steatohepatitis

NBPF7 - Putative neuroblastoma breakpoint family member 7

NSUN2 - RNA cytosine C(5)-methyltransferase NSUN2

NT5C - 5'(3')-deoxyribonucleotidase, cytosolic type

PALLD- Palladin

PDCD5 - Programmed cell death protein 5

PIGR - Polymeric immunoglobulin receptor

PLIN2 - Perilipin-2

PPP1R7 - Protein phosphatase 1 regulatory subunit 7

PRKACB - cAMP-dependent protein kinase catalytic subunit beta

PTMS - Parathymosin

RAB23- Ras-related protein Rab-23

RBM4 - RNA-binding protein 4

RBM5 - RNA-binding protein 5

SH3BGRL3 - SH3 domain-binding glutamic acid-rich-like protein 3

SLC26A2 - Sulfate transporter

SLC35A3 - UDP-N-acetylglucosamine transporter

SNX1 - Sorting nexin-1

SNX6 - Sorting nexin-1

TSN - Translin

TSTA3 - GDP-L-fucose synthase

Abstract

Background & Aims: Non-alcoholic fatty liver disease (NAFLD) is a global epidemic, manifested in over 10% of the world population, although often is undiagnosed. Liver biopsies are the only gold standard available for the confirmation of the disease state. Other non-invasive diagnostics such as ultrasound and MRI are inaccurate or expensive for routine use. Many markers for disease state are available that detect the onset of inflammation and fibrosis from moderate-to-high accuracy. However, there is a major shortcoming in specific biomarkers that can distinguish NAFLD livers from normal in early stages of the onset of disease.

Methods: SWATH-MS based Data-independent acquisition (DIA) strategy is used in the proteomics analysis due to its ability in quantification of low-abundant proteins with high accuracy. Hence a DIA based strategy is used to evaluate the dysregulated proteins in two in vitro models of hepatic steatosis and to compare it with human liver tissue (n=116) showing progressive stages of NAFLD.

Results: More than 2,500 proteins were identified in HepaRG and human hepatocyte model as well as human liver tissue. Within the hepatocyte model, 40 proteins were dysregulated in steatosis. These proteins were screened in liver tissue and 6 common proteins (PLIN2, ANAX1, H2AFY2, APOE, GCFHR, SNX1) showed significant changes in protein expression. In addition, receiver operative curve (ROC) analysis revealed all markers were relevant in the prediction of the disease state with moderate-to-high selectivity and specificity. Especially, PLIN2, ANAX1 and H2AFY2 showed receiver operative area under the curve (ROAUC) ≥ 0.7.

Keywords

NAFLD, DIA, SWATH-MS, Steatosis, Biomarkers

1. Introduction

Nonalcoholic fatty liver disease (NAFLD) is reported to be the leading cause of end stage liver disease and liver transplantation [1]. It affects up to 30% of the adult population, especially in western countries [2]. Moreover, 10-15% of the patients diagnosed with NASH, develop cirrhosis and hepatocellular carcinoma [3]. Despite the recent advancement in diagnostics, liver biopsy is the only gold standard method for confirmation of NAFLD [4].

Histological testing of liver biopsies, along with multiple scoring scales and algorithms, are employed to classify the progression of the disease [5]. However, liver biopsies are invasive, expensive and poorly accepted among the patient population. In addition, owning to the heterogenicity of lipid localization in liver as seen with severity of steatosis, diagnosis from liver biopsies may be biased and lead to misclassification [6]. Noninvasive imaging techniques have improved over time and are used to help determine the progression of steatosis, inflammation, hepatocellular ballooning, fibrosis, as well as liver stiffness [4, 7]. However, ultrasound fail to detect steatosis with the grade of < 20%, in addition, more precise diagnostics such as MRI, are expensive for routine testing [4]. Hence, biomarkers that best predict the severity and progression of disease state are therefore needed for diagnosis and therapy. In addition, it can also greatly improve the current drug development process for NAFLD. Due to the complexity involved in the disease progression of NAFLD, it is highly unlikely that a single biomarker can precisely delineate steatosis and NASH group. Henceforth, a list of proteins that are dysregulated with disease progression alongside other variables (such as age,

gender and diabetes) would present the robustness for probable biomarker identification.

With the advent of novel proteomics technologies, multiple studies have reported to identify various biomarkers in NAFLD. Bell et al. reported 55 as well as 15 proteins that were dysregulated in NAFLD and NASH with advanced fibrosis along with NASH and NASH with advanced fibrosis, respectively [8, 9]. However, not much is known about the dysregulated proteins between simple steatosis and NASH groups. Proteomics employing sequential acquisition of theoretical mass spectra (SWATH-MS) is a powerful tool that aims to determine the relative amount of proteins from low-throughput biological samples [10-12]. Data independent acquisition based total protein approach (DIA-TPA) is more commonly used in untargeted proteomics, to compare the protein levels between the groups. This approach is sensitive to capture the low abundant proteins extracted from cells as well tissue homogenates, and hence an ideal tool for biomarker discovery [12].

The objective of this study is to identify a list of secretory hepatic proteins associated with the progression of disease. To achieve this, we compared the relative protein expression in human hepatocyte and HepaRG-based models for hepatic steatosis with human liver tissue showing progressive stages of NAFLD. This comprehensive comparison provides a panel of proteins that may act as probable biomarkers of disease state prediction.

2. Material and Methods

2.1 Chemicals and Reagents.

TPCK-treated trypsin, trypsin digested β-galactosidase and mass spectrometer tuning solution were purchased from SCIEX, Framingham, MA. Acquity UPLC Peptide BEH C18 analytical column and VanGuard pre-columns were procured from Waters Corp. (Waltham, MA). IN). Bovine serum albumin, sodium deoxycholate and iodoacetamide (IAA) were procured from Sigma Aldrich (St. Louis, MO). MS grade acetonitrile and formic acid were purchased from ThermoFisher, Waltham, MA.

2.2 Cell Culture

The HepaRG and cHH were cultured as per manufacturer's protocol.

Cryopreserved pooled human hepatocytes (cHH). Pooled donors of cHH, were obtained from Xenotech, LLC (Kansas City, KS). Demographic information of the 5 pooled cHH is listed in **Table 4.1**. cHH were thawed in a 37 °C water bath for ~80 seconds. Vials were inverted transferred to pre-warmed OptiThaw media (K8000). Upon centrifugation at 100 x g for 5 minutes at room temperature or 2-8°C, cells were washed. After thawing, the cell pellets were resuspended in OptiPlating media. The cells were diluted to 1 million cells/mL and seeded onto the collagen coated 6-well plate. Following this, the cells were allowed to attach for 4 h and the media was changed to OptiIncubate Hepatocyte Media (K8400) and allowed to incubate for 24 h. Upon 24 h, media containing 0.5 mM of fatty acids (oleate and palmitate 1:2) conjugated to BSA were added to treatment group in comparison to control. After 72 h of treatment, the cells were washed and scraped (CELLTREAT Scientific Products, Catalog No. #229310, Pepperell, MA) in ice-cold DPBS (ThermoFisher Scientific, Catalog No. #14040117, Waltham, MA) and 5-6 million cells from each replicate was pooled for further analysis.

HepaRG cells. Frozen undifferentiated HepaRG cells were purchased from Biopredic International (Rennes, France). The cells were grown for 2 weeks in William's E media (ThermoFisher Scientific, Catalog No. #12551032, Waltham, MA), containing 1% glutamine (ThermoFisher Scientific, Catalog No. # A1286001, Waltham, MA), 1% penicillin/streptomycin and 10% HepaRG growth supplement (Lonza Inc., Catalog No. #ADD711C, Walkersville, MD) in 6-well plate. Following growth phase, cells were differentiated in-house for 2 weeks in William's E media, 1% glutamine, 1% penicillin/streptomycin and 10% HepaRG differentiation supplement (Lonza Inc., Catalog No. #ADD721C, Walkersville, MD) and maintained at 37°C with 21% O2 and 5% CO2 throughout the culture with media changed every 3 days. Upon differentiation, HepaRG cells were cultured in basal media containing William's E media, 100 U/mL penicillin, 100 µg/mL streptomycin, 2 mM glutamine, 5 µg/mL insulin (Sigma-Aldrich, Catalog No. #I-034, St. Louis, MO) and 50 µM hydrocortisone hemisuccinate (Sigma-Aldrich, Catalog No. #H2270, St. Louis, MO). Upon differentiation, the cells were treated with fatty acid media as described above. Approximately 8-10 million cells were harvested by cell scraping for proteomic analysis.

2.3 Oil Red O Staining

Pooled cHH treated with 0.5 mM FA or control for 72h was analyzed for neutral lipid accumulation. Cells were washed twice with 1X PBS and fixed with 4% paraformaldehyde (Sigma-Aldrich, Catalog# 100496, Saint Louis, MO) for 30 min in room temperature (RT). Following this, 500 µl of working Oil Red O (ORO) solution (3 parts of stock ORO and 2 parts of dH2O) (Sigma-Aldrich, Catalog# O0625, Saint Louis, MO) was added and incubated at RT for 20 min. The cells were washed 3X with distilled water (dH2O), and 500 µL of hematoxylin (Sigma-Aldrich, Catalog# H3136, Saint Louis, MO) was added for 1 min to stain the nuclei and their image was obtained using EVOS cells imaging systems

(Life Technologies, Model# M7000, Foster City, CA). The extent of neutral lipid accumulation was quantified after extraction in 100% isopropanol, and the absorbance was measured at 492 nm using a SpectraMax plate reader (Molecular Devices, Model# M2, Sunnyvale, CA).

2.4 Proteomics sample preparation

Cell pellets were homogenized in 500 µl of homogenization buffer (8 M urea, 50 mM Triethylammonium bicarbonate, 10 mM DTT v/v) using a bead homogenizer (Omni Bead Ruptor, Kennesaw GA). 50 mg of human liver tissue (HLT) (n=116) was homogenized in 1000 µl of homogenization buffer, in similar fashion. Samples were then spun at 1,000 g for 5 min and the supernatant was collected. The total protein concentration was determined using Pierce BCA protein assay (ThermoFisher Scientific, Catalog No. #23227, Waltham, MA).

2.5 In-solution trypsin digestion

Protein digestion was conducted as previously described with few adaptations. Samples (250 µg protein) were spiked with 2 µg of BSA and denatured with 25 µL of DTT (100 mM) at 35°C for 30 min in a shaking water bath (100 rpm) (Precision Scientific, Catalog No. #66800, Chicago, IL). After denaturation, samples were alkylated in the dark with 25 µL of IAA (200 mM) for 30 min at room temperature. Samples were then concentrated using the cold water, methanol and chloroform (1:2:1) precipitation method (centrifugation at 10000 rpm, 5 min at 10°C) (Eppendorf, Catalog No. #5810 R, Hamburg, Germany). The protein pellet was washed with ice-cold methanol and suspended in 100 µL of 50 mM ammonium bicarbonate (pH ~8) containing 3% w/v sodium deoxycholate (DOC). Further, TPCK-treated trypsin (10 µg) was added to samples at a ratio of 1:25 (trypsin: protein) and samples were transferred into digestion tubes (PCT MicroTubes,

Pressure Biosciences Inc., Easton, MA). The barocycler was run at 35°C, for 75 cycles with 60-sec pressure-cycle (50-sec high pressure, 10-sec ambient pressure, 25 kpsi). Subsequently, 10 µg trypsin was added again for futher digestion as mentioned above.

Furthermore, to 110 µL of digested peptides sample, 10 µL of acetonitrile (1:1, v/v containing 5% formic acid) was added to precipitate DOC. Samples were spun (10,000 rpm for 5 min at 10°C) to remove the precipitate and 100 µL supernatant was collected. Subsequently, 25 µL of the digested peptide sample was injected on the analytical column and were analyzed using LC-MS/MS method described below.

2.6 LC-MS/MS method and SWATH-MS data analysis

The LC-MS/MS method was used as previously developed without modifications. Homogenate samples were analyzed using SWATH-MS based spectra which were acquired for mass range m/z 400−1100 Da within SWATH window width of 10 m/z resulting in 70 overlapping mass windows per cycle, as described in the previous literature21. The DIA data was searched against Swiss-Prot identifiers (October 2016) and analyzed using Spectronaut™ Pulsar software (version 11.0, Biognosys AG, Schlieren, Switzerland) was used to obtain MS2 signal intensities of fragment ions from DIA data using default settings. Raw data files and search results are available at Japan Proteome Standard Repository (jPOSTrepo JPST000373, ProteomeXchange PXD008594). The SWATH-MS DIA data was matched with the reference spectral library generated from the DDA of pooled human liver samples using Protein Pilot. The DIA data was then normalized and analyzed using “Total protein approach” (TPA), as shown below

$$\text{Protein(i)} = \frac{MS\ signal(i)}{Total\ MS\ signal \times \text{Molecular mass(i)}} \times 10^9 \quad \ldots (1)$$

The absolute protein abundance of each group was represented as pmol/mg of protein.

2.7 Statistical and bioinformatic analysis

Missing data with proteins more than 50% among the samples were omitted from analysis. For cell analysis, the proteins present in all replicates were only considered for analysis. Student's t-test with p value < 0.05 and log2FC cutoff of +/- 0.58 was used to identify differentially expressed proteins in pooled human hepatocytes, and the HepaRG cell lines. The data obtained from the cells were represented as means from quadruples. Proteins from human livers were analyzed using SPSS software v24.0. Protein data from human samples were ln (natural-log) transformed and analyzed using one-way ANOVA and post-hoc analysis for multiple group comparison was conducted using Bonferroni test. In addition, the specificity and selectivity of proteins were analyzed using ROC cures on SPSS. P< 0.05 was considered significant. Prism 8 (GraphPad Inc., La Jolla, CA) was used and SPSS 24.0 (IBM Corp., Armonk, NY) were used for graphing and statistical testing, respectively.

3. Results

3.1 Overall workflow in the identification for dysregulated protein in disease state.

The workflow followed in this manuscript is as illustrated in Fig. 4. 1. The samples obtained from cell model for hepatic steatosis was analyzed along with human liver tissue with progressive forms of NAFLD such as: normal (n=42) and NAFL (n=34) based on Kleiner scoring [5]. The data obtained from spectronaut was filtered for missing data and low abundance proteins which showed highly variable expression pattern. For the cell data, the proteins selected for analysis were present in all replicates (n=4). Similarly, for human tissue (n =116) the set criteria were missing data > 50% among all groupings and CV% < 100% owing to the considerable difference caused by the variables considered. In this

way 2781, 2653, 2580 unique proteins were identified in HLT, cHH and HepaRG cell lines, respectively (Fig. 4. 1). Among the 2653 protein 40 proteins were dysregulated in hepatic steatosis. This was compared with HepaRG steatosis model that showed a similar pattern for 7 proteins as described in Fig. 4. 4. All of these 40 proteins were matched with human liver tissue and using ANOVA and post hoc analysis for multiple comparison Bonferroni test ($p < 0.05$), 6 proteins that showed significant difference were identified (Table 4. 3).

3.2. Effect of fatty acid loading to human hepatocytes.

Upon treatment, the pooled human hepatocytes showed increase in neutral lipid accumulation visualized at 72h utilizing ORO stain (Fig 4.2A and 4.2B). In addition, we noticed a 6-fold increase in absorbance (492 nm) when treated with fatty acids in comparison to the control ($p < 0.01$) (Fig. 4. 2C).

3.3 SWATH-MS based evaluation of proteins in human hepatocytes, HepaRG and human liver tissue.

Upon filtering with set criteria, 2781, 2653, 2580 unique proteins were identified in HLT, cHH and HepaRG cell lines, respectively (Fig. 4.1). Among the groupings, 1967 proteins were common among the three groups (Fig. 4.2D). However, 130 proteins were present in HepaRG and HLT independently showing the abundance of biliary-epithelial-like cells in HepaRG other than the hepatocyte-like cells. In addition, 128 proteins in HLT was not shared in any other groupings due to the presence of cell types than hepatocytes alone.

3.4 Dysregulated protein expression in hepatic steatosis as seen in human hepatocytes.

In this model for hepatic steatosis, 41 proteins were dysregulated using the cutoff for log2FC +/- 0.58 and p-value < 0.05 (Fig.4.3C). Among the 40 proteins, 14 were downregulated (Fig. 4.3A) and 26 were upregulated (Fig.4.3B). Within the 14 downregulated proteins, 5 (PPP1R7, SH3BGRL3, SNX6, NSUN2, PRKACB, RBM4) were

localized in the nucleus and 3 (EIF5B, NT5C, SLC35A3) in the cytoplasm. Secretory proteins, PIGR (-0.66), KRT10 (-0.61) and APOE (-0.61) as well as cytoskeleton proteins CALD1 (-0.95) and PALLD (-0.99) were identified to be significantly downregulated in steatosis. Moreover, among the 26 dysregulated proteins, 6 secretory proteins HMGB1 (0.62), PDCD5 (0.63), APOA4 (0.78), TSTA3 (1.06), ANAX1 (1.31), LTF (2.33). Other proteins involved in lipid storage PLIN2 (2.45) as well as those involved in xenobiotic metabolism GSTA4 (1.13), cytoskeleton remodeling DBNL (0.62) were also upregulated.

3.5 Dysregulated protein expression in hepatic steatosis as seen in combined model of HepaRG and human hepatocytes.

Among the 40 proteins that were dysregulated in human hepatocytes, 7 proteins showed a common signature in HepaRG steatosis model as well. KRT10 and APOE were significantly downregulated. On contrast, TSN, ANXA1, PTMS and PLIN2 showed upregulation to the extent that is similar between hepatocyte and HepaRG model.

3.6 Dysregulated proteins as possible markers of hepatic steatosis.

The 40 proteins that were dysregulated in pooled human hepatocytes were used as a starting point to screen for the protein expression in liver tissue showing differential expression in NAFLD. Among the 40 proteins, 6 showed significant change in expression in disease state (Table 4.3). It was to be noted that, majority of the protein ANXA1, GCHFR, H2AFY2, PLIN2, SNX1 and APOE showed a significant change ($p < 0.05$) in the earlier stages of the disease with steatosis, hepatocellular ballooning with minimal/no lobular inflammation (Fig. 4.5). Moreover, PLIN2 showed an increase with progression of disease from NAFL to NASH (data not shown). Within this pool of 6 proteins, ANAX1 and APOE were secretory proteins and PLIN2 is associated with lipid droplets present in tissue as well as plasma.

The specificity and selectivity of the markers were analyzed using ROC analysis. All 6 markers showed significance, however, PLIN2, H2AFY2 and ANAX1 showed ROAUC of ≥ 0.7 showing the potential to be good markers of disease state (Table 4.4).

4 Discussion

Non-alcoholic fatty liver disease (NAFLD) is the estimated, within next 20 years, the major cause of liver related morbidity and mortality [3]. Liver biopsy is the only confirmatory test for the onset of the disease [1]. Due to the invasiveness, substantial cost as well as variability associated in this testing method, novel non-invasive techniques for the diagnosis of NAFLD have been extensively studied for development, in the recent times. The absence of biomarkers at early stages of disease is responsible for late diagnosis, leading to undiagnosed NAFLD in general population. Therefore, identifying novel biomarkers to detect the onset of the disease as well as disease progression is the need of the hour.

Proteomics is emerging as an efficient tool for its application in exploratory biomarker studies[12]. DIA based SWATH-MS approach is sensitive and highly reproducible method to measure low abundant proteins in complex biological matrices [10, 11, 15]. Hence, SWATH-MS approach was used to screen our samples as illustrated in **Fig. 1.** The global proteome expression, in the invitro HepaRG and hepatocyte model for steatosis was compared with human liver tissue with varying degrees of NAFLD. In addition, the whole cell and tissue homogenate without fractionation was analyzed as it most closely resembled the physiological state.

Upon analyzing the samples, the dysregulated proteins in human hepatocytes model for hepatic steatosis was performed using log2FC and pvalue < 0.05 (Figures 3A–C). A few of the dysregulated proteins observed in human hepatocytes (PIGR, APOE, ANXA, KRT10) are identified in exosomes released from the mouse model of NAFLD,

which validates the prominence of these markers in disease state [16]. Furthermore, our analysis from liver samples from different stages of NAFLD demonstrate that PLIN2, ANXA1, GCHFR, SNX1, H2AFY2 and APOE were significantly altered in NAFLD in comparison to control groups. Moreover, PLIN2 and H2AFY2 showed log fold change (logFC) greater than 0.58 showing more pronounced deregulation. In addition, ROC analysis highlights the selectivity and sensitivity of the markers individually on the disease state prediction **(Table 4)**. It shows that PLIN2, ANXA1 and H2AFY2 are good markers of disease state with ROAUC ≥ 0.7.

It is important to note that most of the markers showed an altered expression pattern in the earlier stages of the disease (NAFL) rather than those observed in NASH (data not shown). This pattern could be because of the bias associated with the comparison of dysregulated proteins from the hepatocyte-steatosis model. Furthermore, the sample size is not large enough to conclude the markers definitively to predict NAFL. Hence, studies with larger sample sizes need to conducted to make more conclusive remarks.

PLIN2 is an abundant hepatic protein associated with lipid droplets. Reports show staining pattern of perilipin to distinguish steatosis and non-alcoholic steatohepatitis in adults and pediatric population [17]. In addition, PLIN2 is concentrated and detected in human urine and is used to differentiate renal carcinomas [18]. However, it is a ubiquitous protein associated with adipocytes differentiation, protumorigenesis and steatosis in cardiomyocytes. Hence, its selectivity for a disease state prediction may lead to low confidence. Therefore, PLIN2 along with other markers needs to be assessed together for understanding NAFLD. ANXA1 is an anti-inflammatory protein secreted by the hepatocytes. Plasma levels of ANXA1 are shown to correlate with fatty liver index in type 2 diabetes patients [19]. Additionally, its role is speculated to attenuate insulin resistance

and hepatosteatosis. The evidences in literature alongside our observation in disease state suggest the probable role AnnexinA1 as a biomarker for NAFL. Nuclear protein H2AFY2, a variant of histone protein H2A is a repressor of gene expression and APOE is a lipoprotein secreted by the hepatocyte that is required for VLDL clearance from the blood. The role of GCHFR and SNX1 is yet to be explored in NAFLD.

Many previous studies have explored the potential of multiple plasma proteins as biomarkers of NAFLD [9, 20]. Most studies, focus on the hallmarks of the disease progression that is inflammation and fibrosis. Thereby, we notice the promising biomarkers in plasma known thus far, such as: adiponectin, CRP, resistin, leptin, RBP-4, IL-6 and TNFα in inflammation and fibrosis; CK-18 fragments in apoptosis; MCP-1 in hepatocellular ballooning; ALT levels in hepatocellular injury; as well as ferritin in fibrosis. However, there is not much know about the early stage predictors of the disease. Hence our work is novel in presenting the dysregulated proteins in earlier stages of NAFLD prior to the onset of NASH.

In summary, a DIA based proteomics using SWATH-MS was implemented to understand the relevant proteins in NAFLD progression. Furthermore, by utilizing in-vitro model for steatosis we were able to delineate the proteins associated particular to the changes in fatty-acid overload. With this approach we streamlined 6 proteins including 2 secretory proteins (ANAX1 and APOE) along with PLIN2 that is all localized ubiquitously with lipid droplets. Our work adds valuable information for future investigations and clinical biomarker research.

5 Figure and Tables

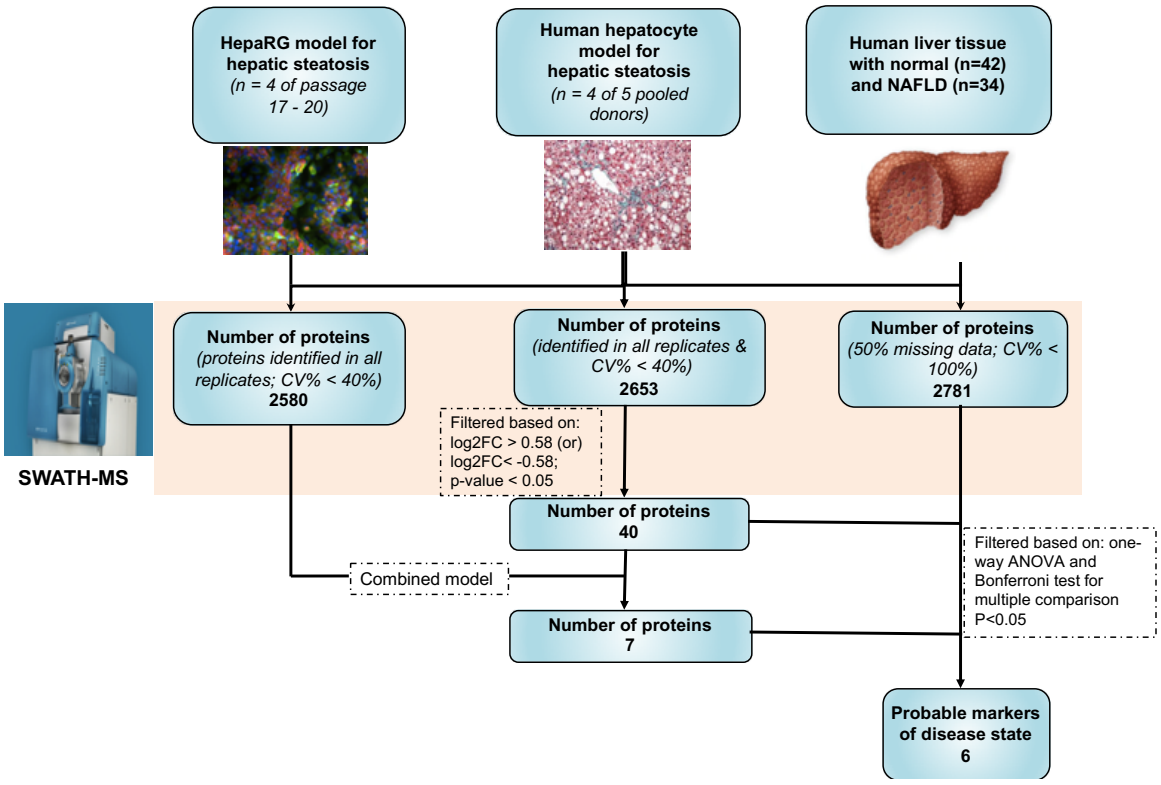

Figure 4.1: Overview of the workflow followed in the biomarker identification.

The HepaRG and pooled human hepatocytes model for hepatic steatosis were collected and analyzed using SWATH-MS. Simultaneously human liver tissue (n=116) homogenates were screened using SWATH-MS and the data was analyzed using spectronaut. Using set criteria, the number of proteins identified in the groups were 2580, 2653 and 2781 in HepaRG, human hepatocytes and human liver tissue, respectively. Finally, the dysregulated from human hepatocytes was matched with human liver tissue to identify 6 markers of disease state.

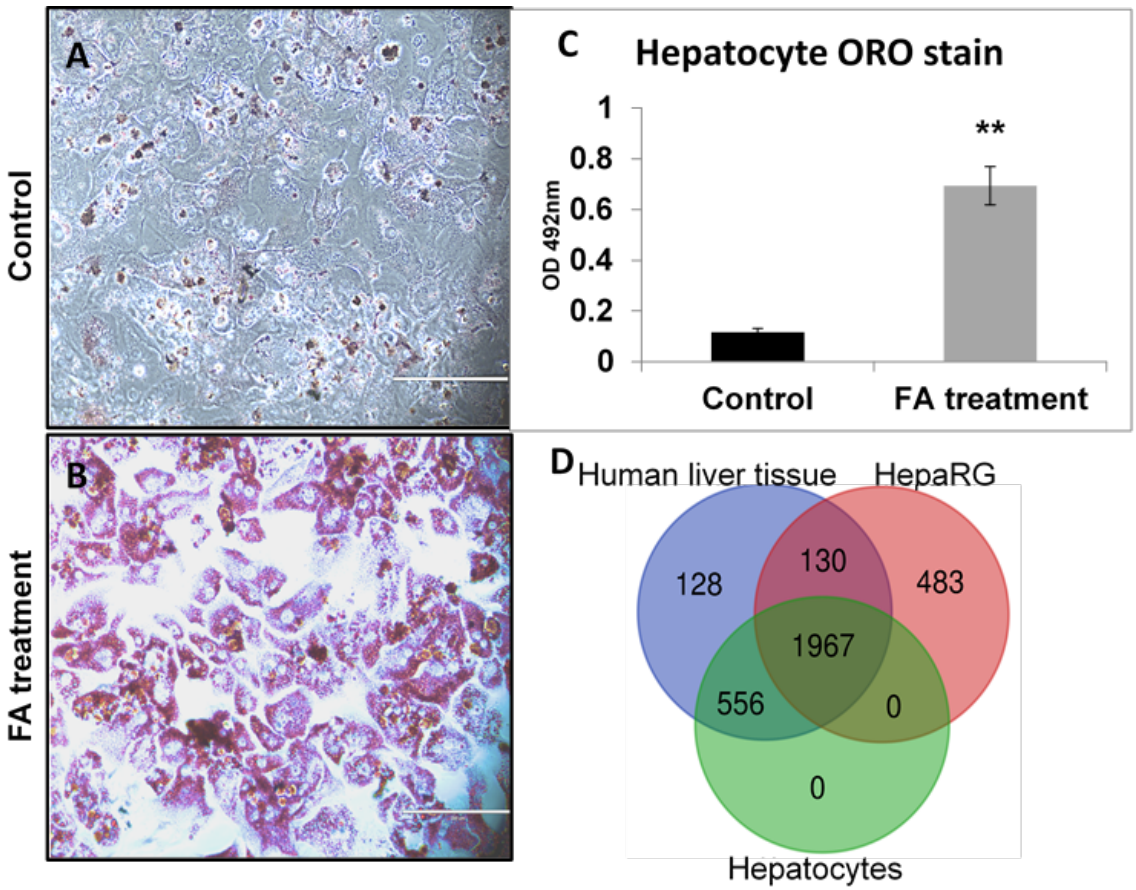

Figure 4.2: Hepatic steatosis in human hepatocytes.

Accumulation of neutral lipids upon the addition of fatty acid mixture to the pooled human hepatocytes is shown in the panel A and B. (C) Shows the 6-fold increase in the absorbance of neutral lipids in the treatment when compared to the control. The results are represented as mean of quadruple with $p < 0.01$. (D) Shows the Venn diagram of the number of identified proteins in human liver tissue, hepatocytes and HepaRG, with 1967 proteins were common among all groups.

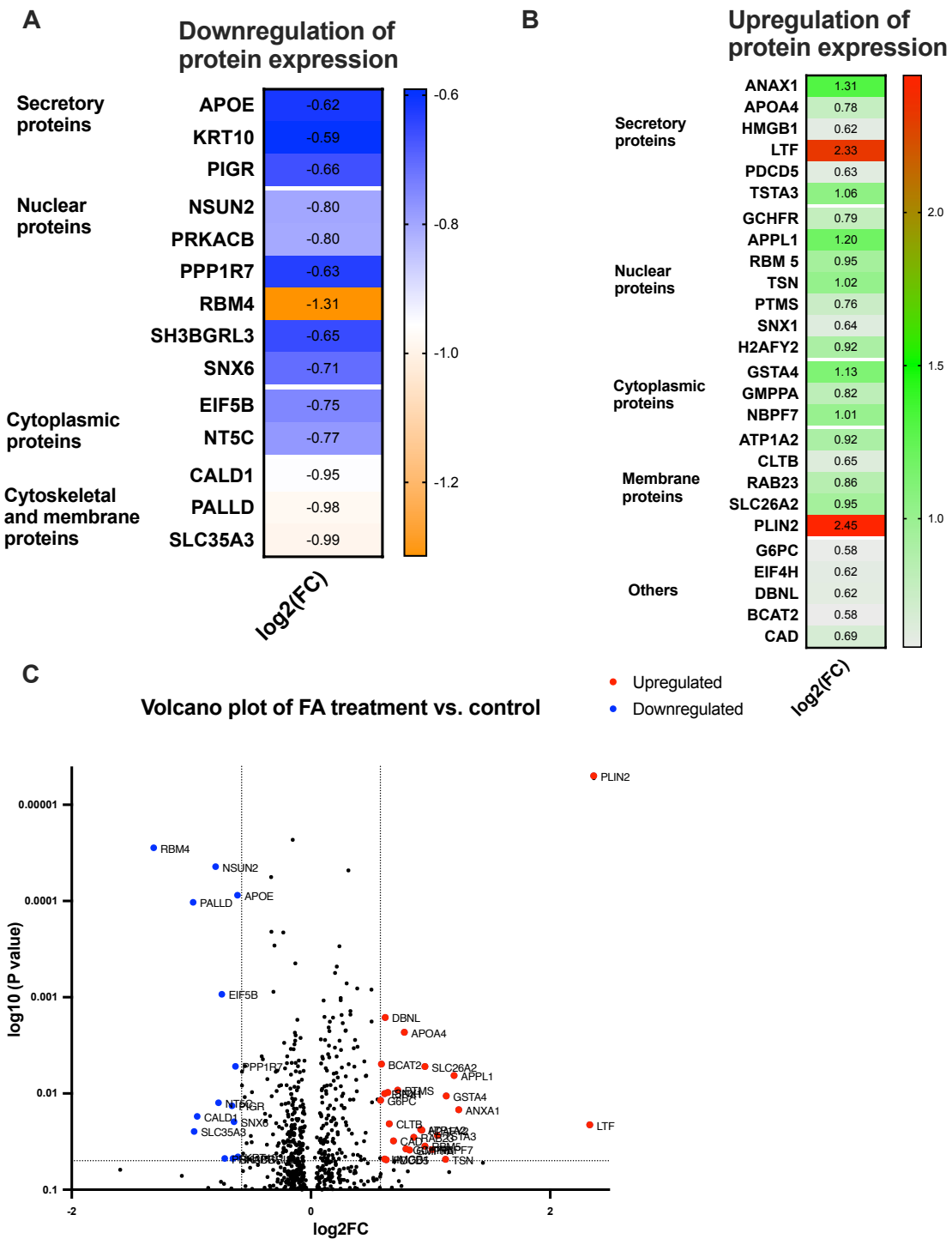

Figure 4.3: Dysregulated protein expression in FA treatment and control in human hepatocytes.

(A) Heat map of downregulated proteins in human hepatocytes (B) Heat map of upregulated proteins in human hepatocytes (C) The volcano plot shows green spots that indicate the upregulated proteins in human hepatocytes (log2FC>0.58 and Pvalue<0.05). The red spots indicate the downregulated proteins in human hepatocytes (log2FC<0.58 and Pvalue<0.05).

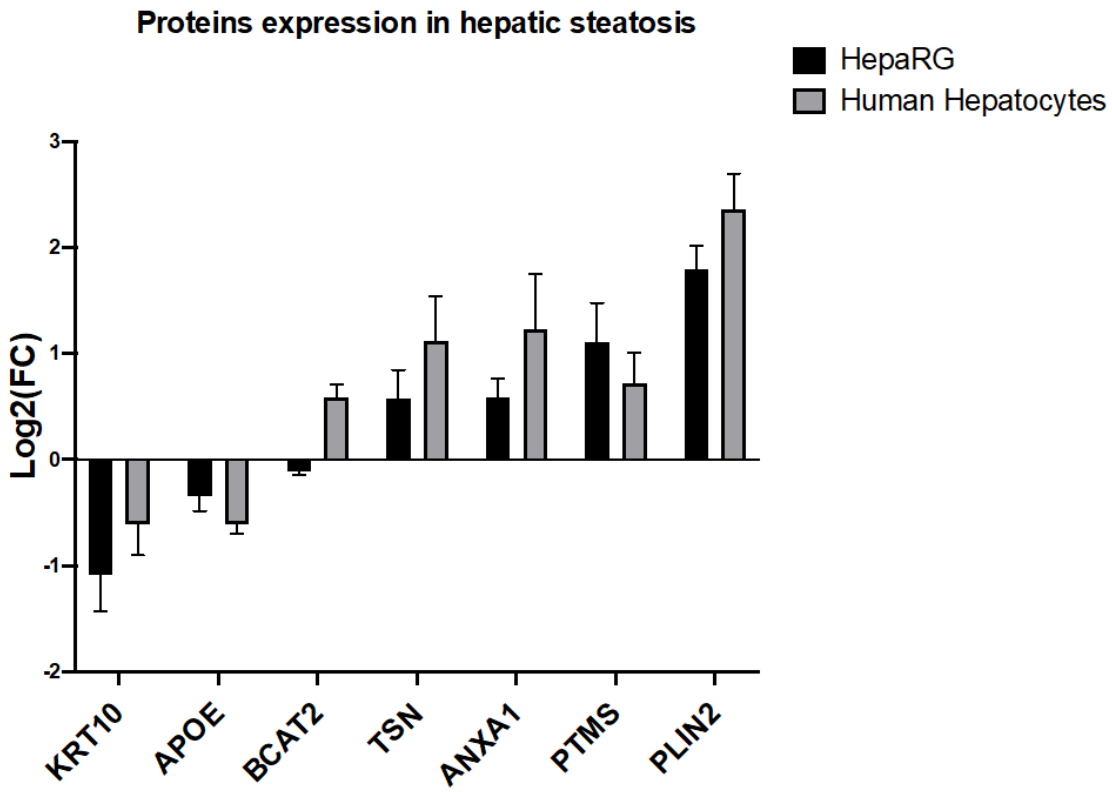

Figure 4.4: Comparison of protein changes in FA treatment and control in HepaRG model and human hepatocytes.

Figure shows proteins that were regulated with p-value <0.05.

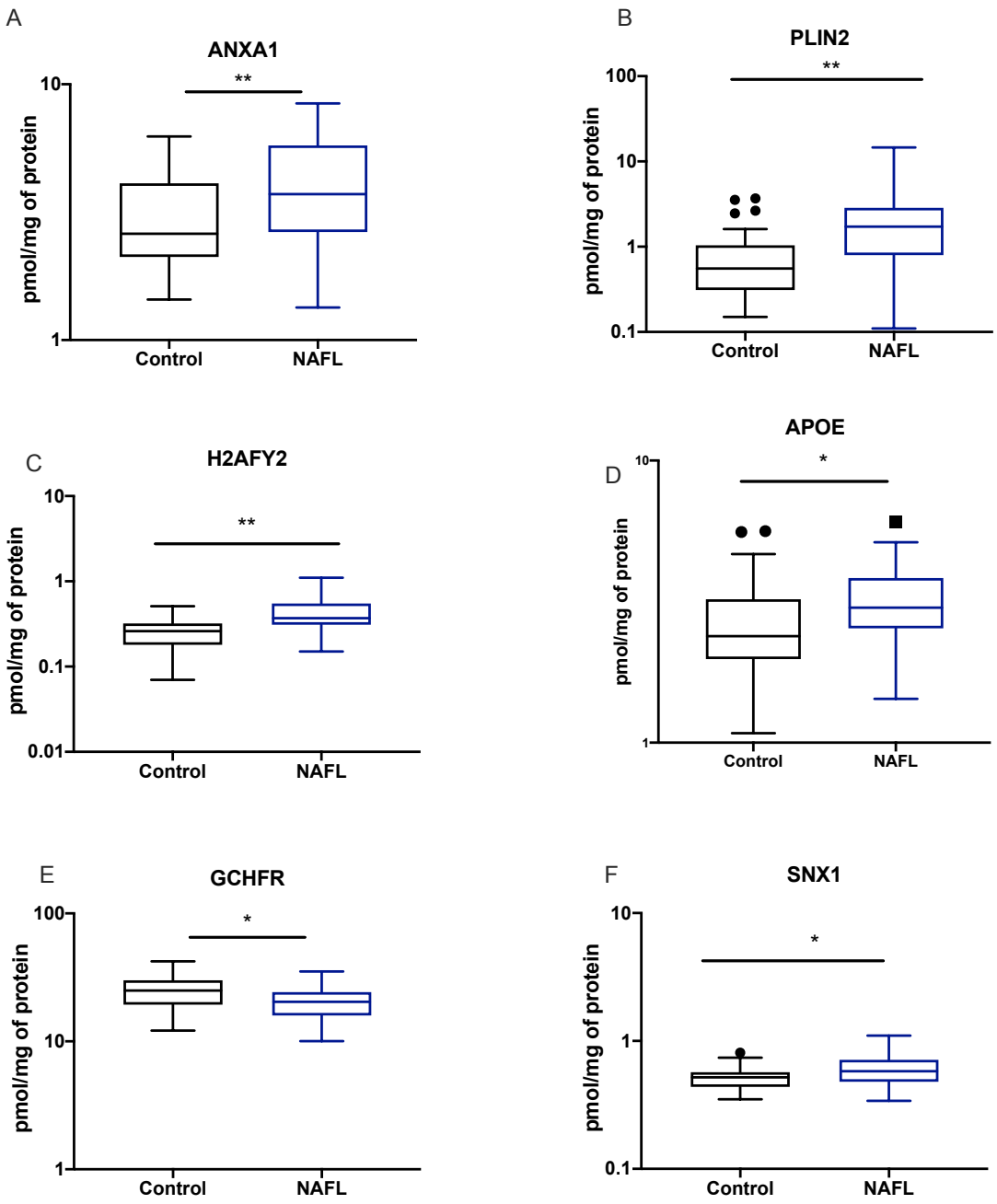

Figure 4.5: Hepatic protein expression shows differential expression in human liver tissue grouped based on control and NAFL.

The differential expression in control (n=42) and NAFL (n=34) in human liver tissue. * denotes P value < 0.05 and ** denotes P-value < 0.01 using Bonferroni multiple comparison on ln-transformation.

Table 4.1 Demographic information of human liver samples

	Control	NAFL
N	42	34
Age (y)	46.5 (41.9 - 51.6)	52.8 (49.5 - 56.4)
Female	16	19
Male	26	15
Caucasian	33	20
Afro-American	8	1
Hispanic	1	3
Body weight (kg)	83.5 (76.2 - 91.5)	89.8 (80.6 - 100.1)
BMI (kg/m^2)	29.3 (26.5 - 32.3)	31.2 (28 - 34.8)

Table 4.2: Demographic information of pooled human hepatocytes.

	Pooled human hepatocyte
N	5
Age (y)	24 - 57 years
Female	2
Male	3
Caucasian	3
Afro-American	1
Asian	1
Cause of death	Head trauma (1); Cerebrovascular accident (3)
Antibody to cytomegalovirus (CMV)	Positive (4); Negative (1)

Table 4.3*:* Data summary of the significantly altered proteins in NAFLD

	Normal		95% CI		NAFLD		95% CI		log2FC
	Mean	N	lower bound	upper bound	Mean	N	lower bound	upper bound	
PLIN2	0.61	42	0.48	0.79	1.51**	35	1.08	2.11	1.29
ANXA1	2.80	42	2.49	3.15	3.81**	35	3.26	4.44	0.44
H2AFY2	0.24	42	0.20	0.27	0.41**	35	0.34	0.48	0.79
SNX1	0.51	42	0.48	0.54	0.58*	35	0.54	0.63	0.20
GCHFR	24.02	42	22.10	26.11	19.48*	35	17.50	21.69	-0.30
APOE	2.48	42	2.21	2.78	3.07*	35	2.74	3.43	0.31

ABBREVIATION: CI - Confidence Interval; FC - Fold Change; N = no. of subjects; PLIN2 - Perilipin 2; ANXA1 - Annexin A1; H2AFY2 - Core Histone; SNX1 - Sorting protein 1; GCHFR - GTP Cyclohydrolase I Feedback Regulator; APOE - Apolipoprotein E. The differential expression in human liver tissue. * denotes P value < 0.05 and ** denotes P-value < 0.01 using Bonferroni multiple comparison on one-way ANOVA

Table 4.4: Receiver Operative Curve (ROC) Analysis for the markers of disease state.

	Area Under Curve (ROAUC)	Std. Error[a]	Asymptotic Sig.[b]
PLIN2	0.770	0.056	0.000
ANXA1	0.700	0.061	0.003
H2AFY2	0.798	0.055	0.001
SNX1	0.669	0.063	0.011
GCHFR	0.688	0.061	0.005
APOE	0.688	0.061	0.005

ABBREVIATION: ROAUC - Reciever operative Area Under the Curve; PLIN2 - Perilipin 2; ANXA1 - Annexin A1; H2AFY2 - Core Histone; SNX1 - Sorting protein 1; GCHFR - GTP Cyclohydrolase I Feedback Regulator; APOE - Apolipoprotein E. The differential expression in human liver tissue. * denotes P value < 0.05 and ** denotes P-value < 0.01 using Bonferroni multiple comparison on one-way ANOVA.

Table 4.5: Dysregulated proteins in hepatic steatosis using a human hepatocyte model.

Localization	Protein Symbol	Protein Name	Molecular process	Log2(FC)
Secretory protein	**ANXA1**	Annexin A1	Immune response	**1.31***
	APOA4	Apolipoprotein A-IV	Cholestrol efflux, lipid homoestatis and lipid transport	**0.78**
	APOE	Apolipoprotein E	APOE vital lipoprotein metabolism	**0.62***
	HMGB1	High mobility group protein B1	Promotes inflamatory response	**0.62**
	KRT10	Keratin, type I cytoskeletal 10	KT10 dysregulation leads to impaired NF-kB activity via inhibition of Akt	**-0.59**
	LTF	Lactotransferrin	Negative regulator of lipopolysccaride mediated signalling pathway	**2.33**
	PDCD5	Programmed cell death protein 5	Apoptosis	**0.63**
	PIGR	Polymeric immunoglobulin receptor	Neutophil degranulation	**-0.66**
	TSTA3	GDP-L-fucose synthase	T cell mediated cytotoxicity	**1.06**
Nuclear proteins	**APPL1**	DCC-interacting protein 13-alpha	Activation of AKT1 pathway in insulin response	**1.21**
	EIF4H	Eukaryotic translation initiation factor 4H	Translation initaition	**0.62**

	GCHFR	GTP Cyclohydrolase I Feedback Regulator	Phenylalanine metabolism and nitric oxide production	**0.65***
	NSUN2	RNA cytosine C(5)-methyltransferase NSUN2	Cell division	**0.81**
	PPP1R7	Protein phosphatase 1 regulatory subunit 7	Positive regulator of protein dephosphorylation	**-0.63**
	PRKACB	cAMP-dependent protein kinase catalytic subunit beta	Cellular response to glycogen synthesis	**-0.82**
	PTMS	Parathymosin	Immune response	**0.76**
	RBM5	RNA-binding protein 5	Apopotoptic pocess	**0.95**
	RBM4	RNA-binding protein 4	RNA binding protein	**-1.31**
	SH3BGRL3	SH3 domain-binding glutamic acid-rich-like protein 3	Cell redox homeostatis	**-0.65**
	SNX1	Sorting nexin-1	Intracellular protein transport	**0.64***
	SNX6	Sorting nexin-1	Intracellular protein transport	**-0.71**
	TSN	Tanslin	Stimulates RNAi production for silencing the transcript from protein expression	**1.02**
	H2AFY2	Core histone macro-H2A.2	Keratinocyte differentiation; regulation of gene expression	**0.92***
Cytoplasmic protein	**GSTA4**	Glutathione S-transferase A4	Xenobiotic metabolism	**1.13**

	GMPPA	Mannose-1-phosphate guanyltransferase alpha	Formation of glycolipids	**0.82**
	NBPF7	Putative neuroblastoma breakpoint family member 7	unknown	**1.01**
	EIF5B	Eukaryotic translation initiation factor 5B	Translational initiation	**0.62**
	NT5C	5'(3')-deoxyribonucleotidase, cytosolic type	Purine catabolism	**-0.77**
Cytoskeletal proteins	**CALD1**	Caldesmon	Increase the stiffness of actin filament and promote cellular contraction	**-0.95**
	PALLD	Palladin	Increase the stiffness of actin filament and promote cellular contraction	**-0.99**
	DBNL	Drebrin-like protein	Reorganization of actin cytoskeleton	**0.62**
Membrane protein	**ATP1A2**	Sodium/potassium-transporting ATPase subunit alpha-2	Sodium potasium transporter involved in wide range of processes	**0.92**
	CLTB	Clathrin light chain B	Translational initiation	**0.65**

	RAB23	Ras-related protein Rab-23	Intracellular protein transport	**0.86**
	SLC26A2	Sulfate transporter	Sulfate transporters involved in bone formation	**0.95**
	PLIN2	Perilipin-2	Lipid storage	**2.45***
Other	**G6PC**	Glucose-6-phosphatase	Converts glucose 6 phosphate to glucose	**0.58**
	BCAT2	Branched-chain-amino-acid aminotransferase	Induce heterogenecity in LD in hepatocytes; BCAA are indicative hepatic stress	**0.58**
	SLC35A3	UDP-N-acetylglucosamine transporter	Carbohydrate transport	**-0.98**

*Footnote - *indicate proteins that showed a significant change in expression in human liver tissue*

6. References

1. Chalasani, N.; Younossi, Z.; Lavine, J. E.; Charlton, M.; Cusi, K.; Rinella, M.; Harrison, S. A.; Brunt, E. M.; Sanyal, A. J., The diagnosis and management of nonalcoholic fatty liver disease: Practice guidance from the American Association for the Study of Liver Diseases. *Hepatology* **2018,** *67* (1), 328-357.

2. Williams, C. D.; Stengel, J.; Asike, M. I.; Torres, D. M.; Shaw, J.; Contreras, M.; Landt, C. L.; Harrison, S. A., Prevalence of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis among a largely middle-aged population utilizing ultrasound and liver biopsy: a prospective study. *Gastroenterology* **2011,** *140* (1), 124-31.

3. Benedict, M.; Zhang, X., Non-alcoholic fatty liver disease: An expanded review. *World J Hepatol* **2017,** *9* (16), 715-732.

4. Pearce, S. G.; Thosani, N. C.; Pan, J. J., Noninvasive biomarkers for the diagnosis of steatohepatitis and advanced fibrosis in NAFLD. *Biomark Res* **2013,** *1* (1), 7.

5. Kleiner, D. E.; Brunt, E. M.; Van Natta, M.; Behling, C.; Contos, M. J.; Cummings, O. W.; Ferrell, L. D.; Liu, Y. C.; Torbenson, M. S.; Unalp-Arida, A.; Yeh, M.; McCullough, A. J.; Sanyal, A. J.; Network, N. S. C. R., Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* **2005,** *41* (6), 1313-21.

6. Vilar-Gomez, E.; Chalasani, N., Non-invasive assessment of non-alcoholic fatty liver disease: Clinical prediction rules and blood-based biomarkers. *J Hepatol* **2018,** *68* (2), 305-315.

7. Bell, L. N.; Theodorakis, J. L.; Vuppalanchi, R.; Saxena, R.; Bemis, K. G.; Wang, M.; Chalasani, N., Serum proteomics and biomarker discovery across the spectrum of nonalcoholic fatty liver disease. *Hepatology* **2010,** *51* (1), 111-20.

8. Niu, L.; Geyer, P. E.; Wewer Albrechtsen, N. J.; Gluud, L. L.; Santos, A.; Doll, S.; Treit, P. V.; Holst, J. J.; Knop, F. K.; Vilsbøll, T.; Junker, A.; Sachs, S.; Stemmer, K.; Müller, T. D.; Tschöp, M. H.; Hofmann, S. M.; Mann, M., Plasma proteome profiling discovers novel proteins associated with non-alcoholic fatty liver disease. *Mol Syst Biol* **2019,** *15* (3), e8793.

9. Jamwal, R.; Barlock, B. J.; Adusumalli, S.; Ogasawara, K.; Simons, B. L.; Akhlaghi, F., Multiplex and Label-Free Relative Quantification Approach for Studying Protein Abundance of Drug Metabolizing Enzymes in Human Liver Microsomes Using SWATH-MS. *J Proteome Res* **2017,** *16* (11), 4134-4143.

10. Jamwal, R.; de la Monte, S. M.; Ogasawara, K.; Adusumalli, S.; Barlock, B. B.; Akhlaghi, F., Nonalcoholic Fatty Liver Disease and Diabetes Are Associated with Decreased CYP3A4 Protein Expression and Activity in Human Liver. *Mol Pharm* **2018,** *15* (7), 2621-2632.

11. Song, Y.; Zhong, L.; Zhou, J.; Lu, M.; Xing, T.; Ma, L.; Shen, J., Data-Independent Acquisition-Based Quantitative Proteomic Analysis Reveals Potential Biomarkers of Kidney Cancer. *Proteomics Clin Appl* **2017,** *11* (11-12).

12. Prasad, B.; Evers, R.; Gupta, A.; Hop, C. E.; Salphati, L.; Shukla, S.; Ambudkar, S. V.; Unadkat, J. D., Interindividual variability in hepatic organic anion-transporting polypeptides and P-glycoprotein (ABCB1) protein expression: quantification by liquid chromatography tandem mass spectroscopy and influence of genotype, age, and sex. *Drug Metab Dispos* **2014,** *42* (1), 78-88.

13. Wiśniewski, J. R.; Vildhede, A.; Norén, A.; Artursson, P., In-depth quantitative analysis and comparison of the human hepatocyte and hepatoma cell line HepG2 proteomes. *J Proteomics* **2016,** *136*, 234-47.

14. Huang, Q.; Yang, L.; Luo, J.; Guo, L.; Wang, Z.; Yang, X.; Jin, W.; Fang, Y.; Ye, J.; Shan, B.; Zhang, Y., SWATH enables precise label-free quantification on proteome scale. *Proteomics* **2015,** *15* (7), 1215-23.

15. Povero, D.; Eguchi, A.; Li, H.; Johnson, C. D.; Papouchado, B. G.; Wree, A.; Messer, K.; Feldstein, A. E., Circulating extracellular vesicles with specific proteome and liver microRNAs are potential biomarkers for liver injury in experimental fatty liver disease. *PLoS One* **2014,** *9* (12), e113651.

16. Pawella, L. M.; Hashani, M.; Eiteneuer, E.; Renner, M.; Bartenschlager, R.; Schirmacher, P.; Straub, B. K., Perilipin discerns chronic from acute hepatocellular steatosis. *J Hepatol* **2014,** *60* (3), 633-42.

17. Morrissey, J. J.; Mellnick, V. M.; Luo, J.; Siegel, M. J.; Figenshau, R. S.; Bhayani, S.; Kharasch, E. D., Evaluation of Urine Aquaporin-1 and Perilipin-2 Concentrations as Biomarkers to Screen for Renal Cell Carcinoma: A Prospective Cohort Study. *JAMA Oncol* **2015,** *1* (2), 204-12.

18. Purvis, G. S. D.; Collino, M.; Loiola, R. A.; Baragetti, A.; Chiazza, F.; Brovelli, M.; Sheikh, M. H.; Collotta, D.; Cento, A.; Mastrocola, R.; Aragno, M.; Cutrin, J. C.; Reutelingsperger, C.; Grigore, L.; Catapano, A. L.; Yaqoob, M. M.; Norata, G. D.; Solito, E.; Thiemermann, C., Identification of AnnexinA1 as an Endogenous Regulator of RhoA, and Its Role in the Pathophysiology and Experimental Therapy of Type-2 Diabetes. *Front Immunol* **2019,** *10*, 571.

19. Ajmera, V.; Perito, E. R.; Bass, N. M.; Terrault, N. A.; Yates, K. P.; Gill, R.; Loomba, R.; Diehl, A. M.; Aouizerat, B. E.; Network, N. C. R., Novel plasma biomarkers associated with liver disease severity in adults with nonalcoholic fatty liver disease. *Hepatology* **2017,** *65* (1), 65-77.

20. Jamwal, Rohitash, "EFFECT OF NON-ALCOHOLIC FATTY LIVER DISEASE (NAFLD) ON HEPATIC DRUG METABOLISM ENZYMES IN HUMAN" (2018). *Open Access Dissertations.* Paper 811.

CONCLUSIONARY REMARKS

This work demonstrates the utility of in vitro models of hepatocytes to aid in the identification of novel biomarkers in hepatic steatosis. In the first part of this work **(Manuscripts II)**, we characterized different cell lines to choose the appropriate human carcinoma cell line for model development in NAFLD. Using this initial assessment, HepaRG cell line was chosen for model development with lipid loading using oleate and palmitate (1:2) at 0.5 mM for 72 h **(Manuscript III)**. With transcriptomics and proteomics technologies we noticed differential expression of numerous transcripts and proteins involved in lipogenesis, lipid transport and hepatocellular injury. In the transcript levels, we identified significant down-regulation of PXR, CAR and HNF4α as well as its target drug metabolizing enzymes CYP1A2, 2B6, 2C8, 2C9,2C19 and 3A4, phase 2 DMEs including UGT1A6 and 2B7, SULT2A1 and 1E1. In the protein level, lower trends of cytochrome P450 were observed, however the difference was not statistically significant. The variability associated with the increasing passage number in HepaRG cell line as well as the changes involved in the biological processes such protein stabilization, or the relative half-life of mRNA in comparison to proteins may have contributed to this variability observed.

In the final part of the dissertation, we adapted the model, developed in HepaRG to human hepatocytes where the fat over-load led to 6-fold increase in neutral lipid accumulation in hepatocytes showing the characteristics of hepatic steatosis model **(Manuscript IV)**. In addition, 40 proteins showed a significant change in human hepatocyte model for steatosis assessed using SWATH-MS. When these proteins were screened in human liver tissue with progressive stages of NAFLD we noticed 6 proteins including 2 secretory protein that have the potency to predict the early stages of disease state.

We believe this work that is a conjunction of in vitro model NAFLD along with the SWATH-MS, DIA based proteomics approach provides vital information for the identification of novel biomarkers to predict the early stages of the disease. We believe this will help in diagnosis of NAFLD in stages prior to NASH that may aid in more effective treatment before the disease had developed to more progressive stages. In addition, we believe these proteins may also have the potential as predictive biomarkers in treatment and therapy.